Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *Human Resources Research Center*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

THE JOHNS HOPKINS
MAR 9 1953
LIBRARY

CONTENTS

Rigidity As a Function of Reversal and Nonreversal Shifts in the Learning of Successive Discriminations: A. H. Buss 75

Response Strength and Conditioned Stimulus Intensity: W. Kessen 82

A Comparison of Two Methods of Producing Experimental Extinction: G. A. Kimble and J. W. Kendall, Jr. 87

The Role of Anxiety in Serial Rote Learning: E. K. Montague 91

Response Strength As a Function of Drive Level and Amount of Drive Reduction: B. A. Campbell and D. Kraeling 97

The Intelligibility of Speeded Speech: W. D. Garvey 102

Intralist-Interlist Relations in Verbal Learning: D. R. Meyer and R. C. Miles 109

Conditioning and Extinction As a Function of Anxiety: K. W. Spence and I. E. Farber 116

Complex Learning and Conditioning As a Function of Anxiety: I. E. Farber and K. W. Spence 120

Object Location in a Complex Perceptual Field: C. W. Eriksen 126

American Psychological Association

Vol. 45 No. 2 February 1953

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas. Priority in publication is given to articles whose authors assume the full cost of publication.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 45, NO. 2 FEBRUARY, 1953

RIGIDITY AS A FUNCTION OF REVERSAL AND NON-REVERSAL SHIFTS IN THE LEARNING OF SUCCESSIVE DISCRIMINATIONS[1]

ARNOLD H. BUSS

Carter Memorial Hospital, Indianapolis

In experiments on rigidity (**3, 5, 11**) *S*s typically learn one set of responses and then are required to shift to another set of responses. Rigidity is defined in terms of failure to shift. There are several different types of shift. The present paper deals with two kinds of shifts in discrimination learning. The first is discrimination reversal ("reversal shift"), which involves a reversal of positive and negative cues. For example, a discrimination is learned with black as the positive stimulus and white as the negative stimulus; then the opposite discrimination is learned, i.e., with white as the positive stimulus and black as the negative stimulus. The second kind of shift to be investigated, "nonreversal shift," involves a change in the stimulus dimension. An example of this shift may be found in the Weigl test (**7**). After *S*s sort stimuli on the basis of form, they are required to switch to color as a basis for sorting.

It has been demonstrated in animal learning experiments that discrimination-reversal learning is slower than learning of the original discrimination (**1, 4, 6, 10, 13**). Nonreversal shift has also been found to interfere with learning of a subsequent discrimination (**8**). Thus, both reversal shift and nonreversal shift seem to retard learning in situations in which infrahuman organisms learn simple discrimination problems. The present experiment investigated the effect of these variables upon rigidity as manifested by human *S*s in a complex discrimination situation.

METHOD

Subjects.—The *S*s were 40 college students from first-year courses in psychology. They were assigned randomly to two groups of 20 each.

Stimuli.—The stimuli, described in detail elsewhere (**2**), were wooden blocks. There were two heights, three shapes, three colors, and three top surface areas. Each block differed from every other block with respect to form, color, height, or area. The same stimulus was never presented twice to an *S*.

Apparatus and procedure.—The apparatus and procedure were the same as those used in a previous study (**2**). A manually operated shutter controlled the exposure time of each stimulus. The *S* depressed two telegraph keys, one marked

[1] Adapted from a dissertation submitted in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the Department of Psychology, Indiana University. The writer is greatly indebted to Dr. W. K. Estes for his generous aid.

Vec and one marked *not a Vec*. The *S* responded by releasing one of the two keys.

Each stimulus was presented for 5 sec. When the shutter closed, *E* said "Right" if *S*'s response was correct, but *E* said nothing if *S*'s response was incorrect. Reinforcement and nonreinforcement were defined in terms of *E*'s saying "Right" or saying nothing.

Experimental design.—There were two groups of *S*s and three series of stimuli. Group I learned a height discrimination in Series 1: all *Vec* responses were reinforced to short blocks and nonreinforced to tall blocks; all *not a Vec* responses were reinforced to tall blocks and nonreinforced to short blocks. Group II learned a form discrimination in Series 1: all *Vec* responses were reinforced to circular blocks and nonreinforced to noncircular blocks; all *not a Vec* responses were reinforced to noncircular blocks and nonreinforced to circular blocks.

In Series 2 both groups were trained on a height discrimination: all *Vec* responses were reinforced to tall blocks and nonreinforced to short blocks. In Series 3 both groups were trained on a color discrimination: all *Vec* responses were reinforced to blue blocks and nonreinforced to non-blue blocks.

In each of the three series both groups were presented with the same stimuli: Series 1—6 tall and 6 short stimuli; Series 2—12 tall and 12 short stimuli; Series 3—9 blue and 9 non-blue stimuli.

It might be expected that in Series 2 both groups would reflect the effects of prior training. Therefore a group with no previous training (Group III) was borrowed from another experiment being run simultaneously (2). The first series of Group III was the same as the first half of Series 2 of Groups I and II.

Only two responses were available to *S:* releasing a key marked *Vec* and releasing a key marked *not a Vec*. The strength of the *Vec* response is inversely proportional to the strength of the *not a Vec* response. Therefore, it is necessary to consider the strength of only one of these responses. The *Vec* response was arbitrarily selected, and hereafter when a response is referred to without designation, it is the *Vec* response. Also, the terms "positive stimuli" and "negative stimuli" are used with reference to the *Vec* response only.

Results

The relative frequency of occurrence of the *Vec* response is used as a measure of its response strength. The frequency data with respect to the height discrimination learned in Series 1 are presented in Table 1. Group I responded more frequently to short blocks than to tall blocks, the difference being significant at the .05 level of confidence. Group II did not respond differentially to the height of the stimuli.

TABLE 1

Mean *Vec* Responses in Series 1 in Relation to Height of Stimuli

Group	Tall		Short		*t*
	Mean	*SD*	Mean	*SD*	
I	2.45	1.53	3.35	1.06	2.41*
II	2.65	1.17	2.80	1.28	.29

* Significant at the .05 level of confidence.

There was an equal number of tall and short blocks in Series 1 (as well as in Series 2 and 3), which makes comparison of relative frequencies straightforward and simple. However, in each of the three series there were more noncircular blocks than circular blocks, and in Series 1 and 2 there were more non-blue blocks, than blue blocks. In order to make comparisons of frequencies meaningful they were weighted. For example, in Series 1 there were 4 circular blocks and 8 noncircular blocks; therefore the response frequency to circular blocks was doubled. There were similar weightings for the form discrimination in Series 2 and 3 and in each half of Series 3 for the color and height discriminations.

TABLE 2

Mean *Vec* Responses in Series 1 in Relation to Form of Stimuli

Group	Circular		Noncircular		*t*
	Mean	*SD*	Mean	*SD*	
I	3.20	2.04	4.20	1.63	1.88
II	6.00	1.79	2.45	1.96	5.55*

* Significant at the .001 level of confidence.

Table 2 presents data on the form discrimination in Series 1. Group II responded more frequently to circular blocks than to noncircular blocks, the difference being significant at the .001 level of confidence. Group I made more responses to noncircular blocks than to circular blocks, but the difference is statistically insignificant.

In Series 2 both groups were trained on a height discrimination, with tall blocks as the positive stimuli. The height-discrimination data for Series 2 of Groups I and II and Series 1 of Group III (the control) are presented in Table 3. Series 2 was divided into halves in order to demonstrate the changes occurring as the series progressed. In the first half of Series 2 Group I responded more frequently to short blocks than to tall blocks, the difference being significant at the .001 level of confidence. Group II also responded more frequently to short blocks than to tall blocks, but the difference is statistically insignificant. Group III, on the other hand, made more responses to tall blocks than to short blocks, the difference being significant at the .02 level of confidence. Thus, in the first half of Series 2 Group I continued to respond on the basis of the height discrimination previously learned (short blocks positive, tall blocks negative). Group II made no height discrimination. Group III, which had no previous training, made a height discrimination with tall blocks as positive stimuli and short blocks as negative stimuli.

TABLE 3

MEAN *Vec* RESPONSES IN EACH HALF OF SERIES 2 OF GROUPS I AND II AND SERIES 1 OF GROUP III IN RELATION TO HEIGHT OF STIMULI

Group	Tall		Short		t
	Mean	*SD*	Mean	*SD*	
First Half					
I	2.70	1.27	3.80	1.25	3.65**
II	2.25	.70	2.60	1.07	1.12
III	3.75	1.30	2.30	1.14	2.68*
Second Half					
I	4.35	1.79	1.85	1.79	3.45**
II	3.40	1.33	3.05	2.04	.76

* Significant at the .02 level of confidence.
** Significant at the .01 level of confidence.

Group III responded more frequently to tall blocks than Groups I and II but less frequently to short blocks. With respect to response frequencies to tall blocks, t tests reveal the following: I vs. II $t = 1.22$, $p > .05$; I vs. III $t = 2.88$, $p < .01$; II vs. III $t = 4.11$, $p < .001$. With respect to short blocks, t tests reveal the following: I vs. II $t = 3.20$, $p < .01$; I vs. III $t = 3.97$, $p < .001$; II vs. III $t = .86$.

Changes were occurring, and these became manifest during the second half of Series 2, as shown in Table 3. Group I responded more frequently to tall blocks than to short blocks, the difference being significant at the .01 level of confidence. Group II responded slightly more frequently to tall blocks than to short blocks, but the difference was statistically insignificant. Thus, Group II failed to shift to the new discrimination, while Group I did shift.

In the second half of the series Group I responded more frequently to tall blocks than Group II and less frequently to short blocks than Group II; the t values for the differences were 1.91 ($p > .05$) and 3.23 ($p < .01$), respectively.

The *Ss* in Group II did not learn the height discrimination in Series 2, but they were not responding randomly. They were making the form

TABLE 4

MEAN *Vec* RESPONSES IN EACH HALF OF SERIES 2 IN RELATION TO FORM OF STIMULI

Group	Circular		Noncircular		*t*
	Mean	*SD*	Mean	*SD*	
First Half					
I	2.90	2.05	5.05	.92	4.17**
II	6.90	1.09	1.40	1.32	9.30**
Second Half					
I	3.71	1.50	3.55	1.07	.32
II	4.97	1.36	2.90	1.38	3.67*

* Significant at the .01 level of confidence.
** Significant at the .001 level of confidence.

discrimination learned previously, as shown in Table 4. In both halves Group II made more responses to circular blocks than to noncircular blocks, the differences being significant at the .001 and .01 levels of confidence, respectively. In the first half Group I made an opposite form discrimination, responding more frequently to noncircular blocks than to circular blocks; the difference was significant at the .001 level of confidence. An inspection of the make-up of Series 1 and 2 revealed no obvious explanation; in each series there were as many tall circles as short circles and as many tall noncircles as short noncircles. In the second half of Series 2, Group I responded about as frequently to circular blocks as to noncircular blocks.

In the first half of Series 2 Group II made more responses to circular blocks than Group I ($t = 14.87$, $p < .001$) and fewer responses to noncircular blocks than Group I ($t = 10.14$, $p < .001$). Similarly, in the second half Group II made more responses to circular blocks than Group I ($t = 2.78$, $p < .01$) and fewer responses to noncircular blocks ($t = 1.67$, $p > .05$).

In Series 3 both Groups were trained on a color discrimination (blue blocks positive, non-blue blocks negative). The data concerning this discrimination are summarized in Table 5. In the first half neither group responded differentially to the color of the stimuli. In the second half both groups made more responses to blue blocks than to non-blue blocks. The difference was significant at the .02 level of confidence for Group I, but statistically insignificant for Group II.

TABLE 5

MEAN *Vec* RESPONSES IN EACH HALF OF SERIES 3 IN RELATION TO COLOR OF STIMULI

Group	Blue		Non-Blue		*t*
	Mean	*SD*	Mean	*SD*	
First Half					
I	2.31	1.65	2.45	1.50	.23
II	2.38	1.18	2.90	1.18	1.50
Second Half					
I	3.00	1.38	1.75	1.21	2.66*
II	2.90	1.18	2.13	1.54	1.65

* Significant at the .02 level of confidence.

However, in both halves and for both blue and non-blue blocks, none of the differences between groups is statistically significant, all *t* values being less than 1.00. Evidently the effects of differential training were no longer being felt.

Since both groups were trained on a height discrimination in Series 2, it might be expected that there would be some carry-over to Series 3. The data concerning the height discrimination in Series 3 are presented in Table 6.

During the first half both groups responded more frequently to tall blocks than to short blocks. The difference was statistically insignificant for Group I but significant at the .01 level of confidence for Group II. This was the first time that Group II showed evidence of a height discrimination. In the second half both groups responded more frequently to tall blocks than to short blocks, but the differences were statistically insignificant. In both halves, for both tall and short blocks, the differences between groups were small and statistically insignificant.

Although the *Ss* in Group II made a height discrimination in the first half of Series 3, they did not give up the form discrimination, as shown in Table 7. Both groups made more responses to circular blocks than to noncircular blocks, but while the difference was significant at the .01 level of confidence for Group II, it was statistically insignificant for Group I. Group II responded more frequently to circular blocks than Group I ($t = 2.30$, $p < .05$) and less frequently to noncircular blocks than Group I ($t = 2.73$, $p < .01$). In the second half the difference between response frequency to circular blocks and response frequency to noncircular blocks was statistically insignificant for both groups. Thus, in the second half of Series 3 Group II finally extinguished the form discrimination learned in Series 1. All differences between groups in the second half were statistically insignificant.

TABLE 6

MEAN *Vec* RESPONSES IN EACH HALF OF SERIES 3 IN RELATION TO HEIGHT OF STIMULI

Group	Tall		Short		t
	Mean	*SD*	Mean	*SD*	
First Half					
I	2.88	1.43	2.00	1.18	1.94
II	3.56	1.27	1.95	1.32	3.46*
Second Half					
I	2.75	1.17	2.13	1.73	1.14
II	2.60	1.71	2.50	1.12	.21

* Significant at the .01 level of confidence.

TABLE 7

MEAN *Vec* RESPONSES IN EACH HALF OF SERIES 3 IN RELATION TO FORM OF STIMULI

Group	Circular		Noncircular		t
	Mean	*SD*	Mean	*SD*	
First Half					
I	3.10	1.94	2.75	1.23	.61
II	4.40	1.62	1.65	1.32	3.12*
Second Half					
I	3.55	2.45	3.15	1.15	.69
II	3.60	2.44	3.55	1.20	.09

* Significant at the .01 level of confidence.

DISCUSSION

Both reversal shift and nonreversal shift retarded subsequent learning. Learning of a reversed discrimination (Group I) was slower than learning of the original discrimination (Group III). In a comparable study McClelland (9) found somewhat different results. In his study human *Ss* first learned to verbalize certain words and to remain silent on others; then they were instructed to reverse this learning. A group having prereversal training learned the reversed task at a significantly faster rate than a group having no previous training. It should be noted that McClelland's *Ss* were instructed to reverse the task,

while in the present study *Ss* were not instructed to reverse the discrimination. Evidently, the effect of instructions is to produce opposite results in fairly comparable situations.

Nonreversal shift (Group II) resulted in slower learning of the subsequent discrimination than reversal shift (Group I). This may be accounted for in terms of a reinforcement-inhibition theory (12) if a principle of partial reinforcement is added. In Series 2 of Group II there was fortuitous partial reinforcement of the form discrimination learned in Series 1. In the first half of Series 2 there were four circular blocks and eight noncircular blocks. Two circular blocks were tall, and two were short; four noncircular blocks were tall, and four were short. Let us assume that an *S* makes a perfect form discrimination in Series 2. He responds *Vec* to all four circular blocks; since a height discrimination was 100% reinforced in Series 2, the *Vec* responses to the two tall circular blocks are reinforced, and the *Vec* responses to the two short noncircular blocks are nonreinforced—50% reinforcement. The *S* responds *not a Vec* to all eight noncircular blocks; this response is reinforced in the presence of the four short noncircular blocks and nonreinforced in the presence of the four tall noncircular blocks—also 50% reinforcement. In the second half of Series 2 two of the five circular blocks were tall (40% reinforcement), and three of the seven noncircular blocks were short (43% reinforcement). Thus, the form discrimination learned by Group II in Series 1 was 50% reinforced in the first half of Series 2 and 42% reinforced in the second half of Series 2.

This partial reinforcement served to maintain the form discrimination in Series 2, retarding learning of the height discrimination by Group II. In Series 2 of Group I there was no reinforcement of the discrimination learned previously; it was 100% nonreinforced. The prereversal discrimination was extinguished, and the new discrimination was learned by Group I by the end of the second half of Series 2.

Summary

This paper investigated the effects of reversal shift and nonreversal shift upon learning of subsequent discriminations. Two groups of *Ss* were trained on three discrimination problems. The stimuli were wooden blocks which differed from each other in height, form, color, or area. One group learned a height discrimination in Series 1 and the reversed height discrimination in Series 2. Another group learned a form discrimination in Series 1 and a height discrimination in Series 2. Both groups learned a color discrimination in Series 3. A control group learned the height discrimination of Series 2 without previous training.

It was found that (*a*) compared with no previous training both reversal shift and nonreversal shift retarded subsequent learning; (*b*) compared with reversal shift, nonreversal shift retarded learning of both subsequent discriminations. The results were accounted for in terms of partial reinforcement and a reinforcement-inhibition theory of discrimination learning.

(Received July 11, 1952)

References

1. Blum, R. A., & Blum, J. S. Factual issues in the "continuity" controversy. *Psychol. Rev.*, 1949, **56**, 33–50.
2. Buss, A. H. Some determinants of rigidity in discrimination reversal learning. *J. exp. Psychol.*, 1952, **44**, 222–227.

3. Cowen, E. L., & Thompson, G. G. Problem solving rigidity and personality structure. *J. abnorm. soc. Psychol.*, 1951, **46**, 165–176.
4. Ehrenfreund, D. An experimental test of the contuinity theory of discrimination learning with pattern vision. *J. comp. physiol. Psychol.*, 1948, **41**, 408–422.
5. Fisher, S. Patterns of personality rigidity and some of their determinants. *Psychol. Monogr.*, 1950, **64**, No. 1 (Whole No. 307).
6. Gatling, F. P. A study of the continuity of the learning process as measured by habit reversal in the rat. *J. comp. physiol. Psychol.*, 1951, **44**, 78–83.
7. Goldstein, K., & Sheerer, M. Abstract and concrete behavior. An experimental study with special tests. *Psychol. Monogr.*, 1941, **33**, No. 2 (Whole No. 239).
8. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
9. McClelland, D. C. Studies in serial verbal discrimination learning. IV. Habit reversal after two degrees of learning. *J. exp. Psychol.*, 1943, **33**, 457–470.
10. McCulloch, T. L., & Pratt, J. G. A study of the pre-solution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, **18**, 271–290.
11. Rokeach, M. Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.*, 1948, **43**, 259–278.
12. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
13. Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *J. exp. Psychol.*, 1945, **35**, 253–266.

Journal of Experimental Psychology
Vol. 45, No. 2, 1953

RESPONSE STRENGTH AND CONDITIONED STIMULUS INTENSITY[1]

WILLIAM KESSEN

Yale University

In *Essentials of Behavior*, Hull (8) postulated a stimulus intensity dynamism effect (V), according to which both habit ($_sH_R$) and reaction potential ($_sE_R$) are functions in part of conditioned stimulus (CS) intensity. A similar position was implicit in *Principles of Behavior* (6). There, Hull assumed that habit strength depends on the strength of *stimulus trace*, which in turn is a function of stimulus intensity. From either of these assumptions, it follows that CS intensity will influence response strength.

Some studies (1, 4, 5, 7, 9, 10) have found that response measures do vary with CS intensity; others (2, 3, 4) have failed to find such an effect. Of particular importance to Hull's theoretical position is the fact that investigations using *extinction* measures have, in general, fallen into the negative category. Since resistance to extinction provides the only uncomplicated measure of habit strength, this suggests that the effect of CS intensity is probably on $_sE_R$ and not on $_sH_R$. Such a generalization is at best tentative, however, because it is drawn from a number of experiments, which differ in many details. What seems to be needed is an attack on the problem in a single experimental situation, using a number of response measures. The two experiments reported in this paper represent an attempt in this direction.

Experiment I

The most obvious prediction from Hull's theory is that CS intensity will have an effect on response strength *during training*. Experiment I tested this hypothesis, using a procedure where every S was trained to several CS intensities. Indirectly, this procedure should also provide a suggestion as to the mechanism through which CS intensity controls behavior. If the effect is on $_sH_R$, the curve relating response strength to CS intensity should be one which first increases and then decreases because of the greater opportunity for the intermediate intensities to benefit from stimulus generalization. But if the effect is solely on $_sE_R$, a monotonically increasing function is predicted.

Method

Apparatus.—The apparatus was similar to that described by Nagaty (11). The training unit was a box 12 in. long, 5½ in. wide and 9 in. high. Three walls and the top were of wood, painted white inside; the fourth wall was of transparent Plexiglas. The floor was a grid of stainless steel bars, 1/16 in. in diameter, about ½ in. apart. In one wooden wall, a wheel of 3¾-in. diameter was set.

The CS was one of a series of lights provided by ordinary frosted house bulbs of 6, 15, 40, 75, 150, and 300 w. Their intensity values (converted to a logarithmic scale to provide more nearly even increments) were .16, .63, 1.19, 1.50, 1.90, and 2.26 log ft.-candles, measured inside the training box. An even and diffuse illumination of the box was obtained by placing a white wooden tunnel between the light source and the box. The light was further diffused by a sheet of translucent frosted glass between the tunnel

[1] The experiments reported in this paper represent part of a dissertation submitted to the Yale University Graduate School in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to thank Dr. Gregory A. Kimble, who directed the research and who made essential contributions to it.

and the box. The frosted glass also prevented *E*'s giving any uncontrolled visual cues to *S*.

The apparatus was wired so that closing two silent mercury switches by an arm mounted to a synchronous motor simultaneously activated the light circuit and started a .1-sec. timer. The closing of a third mercury switch charged the grid. When *S* rotated the wheel, the timer stopped and the shock went off (or did not go on, if *S* responded within the interstimulus interval). The electric shock delivered to the grid was of 90 v. with a fixed resistance of 250,000 ohms in series with the rat. An electric fan ran continuously during the experimental session to mask very slight noises from the apparatus.

Subjects.—The *S*s were 10 male albino rats of the Sprague-Dawley strain, 90–120 days old when used. All *S*s were housed in the experimental room, which was lighted day and night by an overhead source of about .5 ft.-candles. The *S*s were on an *ad libitum* supply of food and water throughout the experiment, and were tamed by being handled. Each *S* was run through the procedure described below in a single session.

Procedure.—The *S* was put in the box for an adaptation period of 2–5 min. before the experiment proper began. At the end of the adaptation period, each of the six lights was turned on for 15 sec. *without shock*, in order to determine *S*'s unconditioned response to the lights. The seventh pretraining trial was a *null trial* in which the mercury switches were thrown by the motor, but neither light nor shock was presented.

Training proper began on the eighth trial, in which, for the first time, light was followed by shock. These trials were run as follows: (*a*) the light and timer were started, (*b*) 5.8 sec. (±.05 sec.) later the grid was charged, and (*c*) about 5 sec. after *S* responded (whether shock was avoided or not), *E* turned off the light. This last procedure was used to reduce the probability that light reduction would serve as reinforcement to the response. Insofar as possible, trials were spaced from 40 to 100 sec. apart. Because *S*s turned the wheel for longer periods early in training, there was a tendency for these early trials to be farther apart.

Each *S* received 66 training trials. A series was devised such that, after six preliminary trials, each light was followed and preceded by each other light twice. The first light was randomly chosen for each *S*. Starting with the first light-shock trial, 11 groups of six trials were run. In each group, each of the six lights was presented only once, and groups of such trials were separated by a null trial which provided a measure of response strength with light and shock absent. *Every S had 11 trials with every light as conditioned stimulus.* On every trial, *S* could avoid the shock by rotating the wheel within 5.8 sec. of the onset of the light.

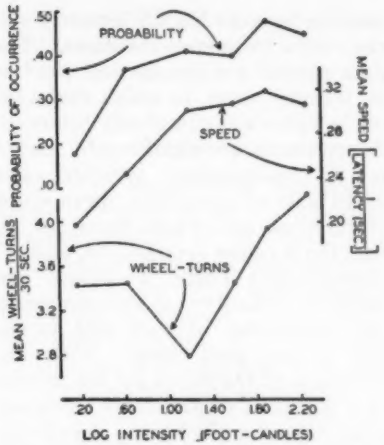

FIG. 1. Response strength during training as a function of log CS intensity, Exp. I. Points represent means based on performance during 60 training trials.

Response measures.—The following data were collected: (*a*) Latency of response measured from the onset of the CS. To make the latency distribution more nearly normal, a reciprocal transformation of latencies was made, giving a speed of response (1/latency) measure. (*b*) Number of avoidance responses, that is, responses with latencies of less than 5.8 sec. This measure was transformed to a probability of occurrence measure by dividing frequency of CR's by number of trials. (*c*) Number of wheel turns made between trials. Because of the variation in intertrial time from rat to rat and from early in training to late in training, the wheel-turn measure was divided by length of intertrial interval (in half minutes) to give a wheel turn per unit time measure.

Results

The *S*s learned the avoidance response. The probability of occurrence of a response with a latency less than 5.8 sec. rose from .02 in pretraining to .13 during the first six light-shock trials to an asymptote of .42 on approximately Trial 40.

The prediction of a relation between CS intensity and response strength was confirmed. Figure 1 shows the

relation between log CS intensity and the three response measures. The data plotted are averages for the last 60 training trials, in which the order of the lights was completely balanced. All measures are significantly related to the independent variable. For probability of occurrence, mean speed of response and wheel turns per 30 sec., the p values are .005, .01, and .05 respectively. These p values are based on an analysis of variance using the interaction of rats and intensity as the error term. With the exception of the curve for the wheel-turn measure, which shows an inversion, the functions are negatively accelerated and increasing.

Experiment II

Experiment I showed that CS intensity is significantly related to response strength *during training*. Moreover, the shapes of the curves suggest that the effect is probably more on ${}_{S}E_{R}$ than on ${}_{S}H_{R}$. A more direct test of this notion would be provided by training different groups of *S*s at different CS intensities and then extinguishing them at several intensities. If the stimulus intensity effect is on ${}_{S}H_{R}$, the number of responses elicited during *extinction* should be significantly related to the intensity of the training stimulus. If the effect is on ${}_{S}E_{R}$, no such relationship is predicted. Experiment II was designed to provide this test, and further, to check the results of Exp. I when an independent group design is used.

Method

Apparatus.—The apparatus was the same as that used in Exp. I. Only four lights were used as CS; those of 6, 15, 40 and 150 w. (.16, .63, 1.19, and 1.90 log ft.-candles).

Subjects.—The *S*s were 32 male albino rats of the Sprague-Dawley strain, about 150 days old when used. Their maintenance routine was the same as that in Exp. I. Each *S* was run in a single session.

Experimental design.—The experimental design was a 4 × 4 factorial. One marginal represented the four light intensities used during training; the other marginal, the same four intensities used during extinction. The *S*s were assigned at random to the 16 conditions in this design. Thus two *S*s were run for each combination of the four training intensities and the four extinction intensities. When response measures during training only or extinction only are examined, this design collapses to a four-column (intensity during training or extinction) design with eight *S*s in each column. Unlike Exp. I, each *S* was trained at only one intensity.

Procedure.—In most details, the procedure in this experiment was the same as in Exp. I. The *S* remained in the experimental box for a 2-min. adaptation period before training began. Pretraining again consisted of six 15-sec. presentations of light without shock, but in this case only the light to be used as CS was presented. The seventh pretraining trial was a null trial. In this experiment, the number of training trials was shortened because Exp. I showed that *S*s reach an asymptote of performance at about Trial 40. The *S*s in this experiment received 42 avoidance training trials and 7 null trials, in an order such that blocks of six trials were preceded and followed by a single null trial. Whenever *S* responded within 5.8 sec. of the onset of the light, the shock did not come on. At the end of training, the interstimulus time was reduced to 2.0 sec. for one trial, to insure that every *S* received a shock immediately before extinction began.

In extinction, *S*s received 30 trials without shock, and five null trials in a series such that groups of six nonshock trials were separated by a single null trial. If *S* did not respond to the light within 15.0 sec., *E* turned it off ending that trial.

All data recorded in Exp. I were recorded in this study for training and for extinction.

Results

When performance measures during the 42-trial training series are examined, a clear relation to CS intensity is seen. Figure 2 shows the relation for the three response measures. When tested by analysis of variance, all measures are significantly related to CS intensity at the .025 level of confidence or better. With the exception of the difference between the two

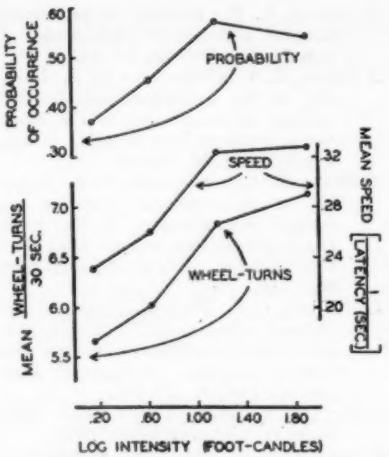

FIG. 2. Response strength during training as a function of log CS intensity, Exp. II. Points represent means based on performance during 42 training trials.

highest intensities for the CR measure, there are no reversals. The reversal in the CR function and the negative acceleration of the other two suggest that the relations are approaching asymptotes. Such an interpretation is supported by the fact that the differences between the groups conditioned to the two highest lights are not significant for any response measure. All other differences (15 comparisons) are significant at the .05 level or better, except for the wheel-turn measure on two comparisons of adjacent lights.

Extinction measures were studied to discover whether they showed differences as a function of either training or extinction intensity. The mean numbers of extinction responses (i.e., responses within 15.0 sec. of CS onset) for *Ss* ***trained*** at the four different lights were 11.25, 14.00, 20.25, and 13.12, in increasing order of light intensity. Mean numbers of extinction responses for *Ss* ***extinguished*** at the four intensities were 12.62, 20.00, 12.50, and 13.50. But an analysis of variance shows no reliable differences along either the training or extinction marginals.

Discussion

The two experiments reported in this paper were successful in demonstrating an effect of CS intensity on response strength during training. For this reason they both support Hull's (**8**) hypothesis of a stimulus intensity dynamism effect on sE_R. Neither experiment, however, provides any evidence for an influence on habit. In Exp. I, the curves relating the three response measures to CS intensity failed to show the bowed shape which such an assumption would predict. In Exp. II, the extinction measures showed no significant effect assignable to the intensity of either the training light or the extinction light.

In view of the fact that a regular relationship between CS intensity and response strength was obtained in *training*, the failure of the intensity of the *extinction* lights to influence behavior is surprising. Presumably it is related to the one important difference between training and extinction. In the latter case, shock is omitted. Under such circumstances total drive is reduced considerably. And it may be that responses antagonistic to wheel turning (e.g., "freezing" to the light) become relatively stronger than the CR and obscure the regularity observed in training.

Summary

Two experiments were run to test the hypothesis that CS intensity is directly related to response strength. In Exp. I, ten albino rats were trained to avoid electric shock by responding to one of a series of six lights of differ-

ent intensities. Every light was used as CS, in a balanced order, for every *S*. Three measures of response strength (probability of occurrence, speed of response, and response amplitude) during training were significantly related to CS intensity ($p = .05$ or better).

In Exp. II, 32 rats were run in the same avoidance situation, with four intensities of CS. Each of four groups of eight *S*s each was trained to a different intensity. Each training group was then split, and two *S*s from each training group were extinguished at each intensity. Performance measures (the same as in Exp. I) *during training* were significantly related to CS intensity ($p = .025$ or better), but extinction measures were related neither to intensity of training lights nor to intensity of extinction lights.

In general, these results support the postulation of a stimulus intensity effect on response strength but suggest that the effect is on performance rather than learning.

(Received July 21, 1952)

References

1. Berlyne, D. E. Stimulus intensity and attention in relation to learning theory. *Quart. J. exp. Psychol.*, 1950, **2**, 71–75.
2. Carter, L. F. Intensity of conditioned stimulus and rate of conditioning. *J. exp. Psychol.*, 1941, **28**, 481–490.
3. Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: I. The conditioned eyelid response to light. *J. exp. Psychol.*, 1948, **38**, 690–696.
4. Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. *J. exp. Psychol.*, 1949, **39**, 35–40.
5. Hovland, C. I. The generalization of conditioned responses. II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291.
6. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
7. Hull, C. L. Stimulus intensity dynamism (*V*) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67–76.
8. Hull, C. L. *Essentials of behavior*. New Haven: Yale Univer. Press, 1951.
9. Kupalov, P. S., & Gantt, W. H. The relationship between the strength of the conditioned stimulus and the size of the resulting reflex. *Brain*, 1927, **50**, 44–52.
10. Kupalov, P. S., Lyman, R. S., & Lukov, B. N. The relationship between the intensity of tone-stimulus and the size of the resulting conditioned reflexes. *Brain*, 1931, **54**, 85–98.
11. Nagaty, M. O. The effect of reinforcement on closely following S–R connections: I. The effect of a backward conditioning procedure on the extinction of conditioned avoidance. *J. exp. Psychol.*, 1951, **42**, 239–246.

Journal of Experimental Psychology
Vol. 45, No. 2, 1953

A COMPARISON OF TWO METHODS OF PRODUCING EXPERIMENTAL EXTINCTION[1]

GREGORY A. KIMBLE
Duke University

AND JOHN W. KENDALL, JR.
Yale University

The experiment reported in this paper compares two methods of producing experimental extinction. One method is that which Guthrie (1) calls the "toleration" method and recommends for the breaking of undesirable avoidance habits. In this method the signal for the avoidance reaction is introduced gradually by increasing its intensity slowly. Under optimal conditions, according to Guthrie, it should be possible to extinguish a response by this technique without the overt occurrence of the response. The other method is the conventional one of presenting the conditioned stimulus (CS) several times without reinforcement. Guthrie sometimes calls this the "exhaustion" method, and predicts that the extinction of a response in this way will require many more extinction responses than the toleration procedure.

As Sheffield (6) has pointed out, the most obvious deduction from Hull's (2) theory would predict the *same* number of responses to extinction under the two procedures. For this reason, he feels that the present experiment is a critical test of Hull's and Guthrie's theories of extinction. Although we do not believe that a single, critical test of the two theories is possible, we agree with Sheffield that an empirical attack on the problem is long overdue, and that results favoring Guthrie's position would necessitate a reconsideration of Hull's explanation of extinction.

[1] This paper is a report of the junior author's senior essay, done at Yale University under the senior author's direction. We wish to thank Prof. Neal E. Miller for a critical reading of the manuscript.

METHOD

Apparatus.—The apparatus used in this investigation is the same as that described by Kessen (5). It consists of a training box with an electrifiable grid as the floor. One side of the box is of transparent Plexiglas through which diffuse illumination can be introduced as a CS. Into one end of the box is inserted a movable wheel which can be rotated to turn off the shock used as the unconditioned stimulus (UCS).

For this experiment, the CS was provided by a 75-w. bulb which was operated on 70 v., and gave a training intensity of 20 ft.-candles. The light circuit was wired through a variable transformer so that the intensity of the CS could be raised from zero to the training intensity by manipulating voltage. Operating the light on 70 v. rather than 110 v. made this procedure somewhat easier and also minimized color changes when the CS was presented at several intensities.

The shock used as the UCS was of 90 v. with a fixed resistance of 250,000 ohms in series with the rat.

Subjects.—Twenty male albino rats of the Sprague-Dawley strain, 100–150 days old, were *S*s in the experiment. They were housed in the experimental room, kept on an ad libitum supply of food and water and were tamed by being handled.

Procedure.—Pretraining, training, and extinction were completed in a single session. Each *S* was placed in the training box for a 2-min. adaptation period and then given a pretraining series of five 15-sec. presentations of the CS without reinforcement. Following pretraining, each *S* was given 60 trials of avoidance training separated by a minimum of 40 sec. On each such trial the CS and UCS were turned on in succession with a 5.8-sec. interstimulus interval. By rotating the wheel, *S* could turn off the shock, or avoid it by making the same response

during the 5.8-sec. interstimulus interval. The CS was turned off about 5 sec. after *S*'s response, whether shock was avoided or not. Finally, *Ss* were extinguished by either the toleration method or the exhaustion method. The assignment of rats to extinction condition was done at random before beginning the experiment.

The *Ss* in the first condition were given seven toleration "trials" immediately. On each such "trial," the CS intensity was raised in 5-v. steps from some lower value to the training level of 20 ft.-candles. Each step lasted 15 sec. and a response was recorded if it occurred in any of these 15-sec. periods. Successive "trials" started at progressively higher voltages, in order to adapt *S* to the sudden onset of stimulation in case this variable was more important than intensity per se. The first toleration "trial" started at an intensity of zero; the second started at 10 v., the third at 20 v., the fourth at 30 v., and so on, until on the eighth trial the CS was turned on at the training intensity. Counting from this eighth trial, the *S* was then given 30 exhaustion trials.

The *Ss* in the exhaustion group rested in the apparatus for 16 min. (the time required for the toleration procedure in the other group) and then received 30 ordinary extinction trials, in which the CS was presented without shock for 15 sec. on each trial and a response counted if it occurred in this period.

A third group of 32 *Ss* from the same population, conditioned by Kessen (5) in the same apparatus and extinguished *immediately*, provided data in terms of which the effect of the 16-min. rest may be assessed.

Results

The data most directly pertinent to a comparison of extinction by the toleration and exhaustion methods are mean numbers of extinction responses. For the toleration method this value is 6.5; for the exhaustion method it is 14.4 ($p = .01$). The average of 6.5 responses obtained in the toleration group is made up of 5.4 responses which occurred in the toleration procedure and 1.1 responses which occurred in the subsequent exhaustion series. But only two *Ss* in the toleration group gave *any* responses in the exhaustion test. One *S* gave nine responses; the other gave two. It is quite clear that, as Guthrie predicts, the toleration procedure results in very efficient elimination of an avoidance reaction.

To make this point as convincingly as possible, the following facts are cited: (*a*) During training the two groups were similar. Mean number of CR's during training were 36.0 for the toleration group, and 38.6 for the exhaustion group ($p = .40$). (*b*) The introduction of the 16-min. rest for the exhaustion group had no significant effect on their extinction behavior. Kessen's 32 animals, which were given 42 trials in the same apparatus, gave an average of 14.7 responses in 30 extinction trials. This value is not significantly different from the measure reported above for our exhaustion group. (*c*) This difference in number of extinction responses occurs in spite of the fact that our procedure gave the *Ss* in the toleration group 92 chances to respond as against 30 in the exhaustion group. (*d*) Such CR's as occurred in the toleration procedure were, on the average, separated in time by a greater amount than those which occurred in the exhaustion procedure. This difference in the degree of massing favors a result opposite to the one obtained. (*e*) The few responses obtained from our toleration group in the exhaustion series all occurred in the first 15 extinction trials. It seems unlikely that the toleration procedure merely produced a temporary suppression of the response.

Discussion

This demonstration of the effectiveness of the toleration procedure in producing experimental extinction calls into question the mechanism proposed by Hull (2) to account for extinction. According to Hull (2), extinction results from the accumulation of an inhibitory potential ($\dot{I}_R$)

which subtracts from reaction potential and eventually leads to the elimination of the response. Furthermore, according to Hull, the development of I_R requires the occurrence of the response. This theory leads to the prediction that our two groups should have produced the same number of responses in the extinction series. Our results, therefore, constitute a serious disconfirmation of this portion of Hull's theory.

This last point does not mean that we believe that Hull's position cannot handle the results of this experiment. An explanation of them which uses ideas implicit in *Principles of Behavior* (2) is developed briefly in the paragraphs which follow.[2]

Although this fact seems never to have been recognized by Hull, his theory contains an alternative account of extinction in which the phenomenon is explained as a reduction of habit strength ($_sH_R$). The equations which relate habit strength to the amount of reinforcement are such that a change in the amount of reinforcement will result in a change in the maximum of $_sH_R$. Although Hull demonstrates that an increase in the amount of reward will lead to an increase in the maximum of $_sH_R$, he apparently did not see that a *decrease* in the amount of reward would produce a decrease in habit, or that, in the special case of the reduction of amount of reward to zero (extinction), $_sH_R$ would gradually reduce to zero. The concept of inhibition, by this argument, is expendable, except for the explanation of the effects of massed practice and the phenomenon of spontaneous recovery. But the position still seems to demand the occurrence of a response for extinction to occur. This latter consideration brings up the next point in our explanation.

Avoidance reactions in avoidance training experiments are mediated responses. That is, they depend for their elicitation upon an acquired drive of fear. Assuming that fear is an internal *response*, it becomes obvious that, if this response can be extinguished, behavior depending on it will automatically be extinguished. We believe that this is what happens in extinction by the toleration method. To describe the process as we think it occurs, we shall need two further concepts from Hull's theory: generalization decrement and the generalization of extinction.

In extinction by the toleration procedure, the stimulus for a fear reaction is presented first in a highly generalized form. This stimulus presumably tends to call out the conditioned fear, but to so small a degree that the instrumental response does not occur. Not because of the accumulation of inhibition, but simply because it is not reinforced, the fear reaction is partially extinguished and this extinction generalizes to adjacent stimulus intensities. Thus, when the intensity of the CS is raised slightly, the fear reaction will occur again, but still in a weak form, and the process described above is repeated. Successive application of the fear stimulus at intensities too weak to produce the avoidance reaction may then finally eliminate it by extinguishing its acquired motivational basis.

Summary

Two groups of male albino rats were trained in a wheel-turning situation to avoid a shock and were then extinguished by two different procedures. One group was subjected to extinction by a toleration method in which the conditioned stimulus is introduced gradually. The other group was extinguished by the conventional procedure of presenting the conditioned stimulus at full intensity without the shock. Mean numbers of responses in the extinction series were 6.5 for the toleration group, and 14.4 for the other group ($p = .01$). This result confirms one of Guthrie's theoretical ideas and necessitates a reformulation of Hull's theory of extinction. Such a reformation, based on the idea that the toleration procedure extinguishes a mediating fear reaction, was attempted.

(Received July 21, 1952)

[2] For several reasons, we regard *Principles of Behavior* as a better exposition of theory than Hull's more recent book (3), and this account is based entirely on the former work. With certain extensions, like the one recently worked out by Jones (4), the newer theory will handle the present results. But the development of these extensions is beyond the scope of this paper.

References

1. Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.
2. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
3. Hull, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1951.
4. Jones, M. B. Effect, change, and expectation of reward. *Psychol. Rev.*, 1952, **59**, 227–233.
5. Kessen, W. Response strength and conditioned stimulus intensity. *J. exp. Psychol.*, 1953, **45**, 82–86.
6. Sheffield, F. D. Hilgard's critique of Guthrie. *Psychol. Rev.*, 1949, **56**, 284–291.

Journal of Experimental Psychology
Vol. 45, No. 2, 1953

THE ROLE OF ANXIETY IN SERIAL ROTE LEARNING[1]

ERNEST K. MONTAGUE
Brooke Army Hospital

A number of investigators (**1, 2, 6, 7**) have conceived of anxiety in the laboratory animal as a not directly observable, acquired, anticipatory response to pain. Anxiety, so defined, is assumed to have at least two of the functional properties of primary drives as these are treated in the Hull-Spence (**4, 9**) formulation of the roles of drive and drive reduction, i.e., its reduction should be reinforcing or should strengthen responses immediately preceding its reduction, and it should intensify any response tendencies that are present during its period of evocation. Within this framework, the strength of response tendencies is conceived to be a multiplicative function of their habit strengths and the general drive level of the responding organism. While the above investigators have been primarily concerned with drive reduction, this paper is more generally concerned with the intensifying or augmenting aspects of anxiety, as drive, on response tendencies of human *Ss*.

Taylor (**10**) has found evidence for this augmenting role of anxiety in an eyelid conditioning experiment with human *Ss*. Defining anxiety in terms of responses to a questionnaire, she found significantly faster conditioning in her anxious *Ss* than in her nonanxious *Ss*. Other studies, Welch and Kubis (**11, 12**) and Rosenbaum (**8**), give added evidence of heightened response tendencies in the presence of anxiety. However, certain studies in verbal learning show quite opposite results. Malmo and Amsel (**5**) found slower serial verbal learning with anxious *Ss* than with nonanxious controls. Montague and Schneider, in an unpublished study, used 20 *Ss* from each of Taylor's original conditioning groups and found slower rote serial learning of nonsense syllables for the anxious *Ss* than for the nonanxious. Although these results appear to be inconsistent with those of Taylor, certain considerations relating to the concept of drive indicate that the effect of anxiety on performance may well vary with the nature of the task involved.

First might be considered the presence of internal and external responses which are connected to the drive stimulus of anxiety. In classical conditioning it may be assumed that these responses present in high anxiety are not incompatible with the single response being learned and indeed may be facilitating. In verbal learning, however, these response tendencies may be much less compatible with the responses necessary for successful performance. Second is the factor of the relative strength and number of correct tendencies as compared with those of the incorrect tendencies in the learning situation. If wrong responses are stronger and more frequently elicited initially in a particular situation, added drive may differentially augment these incorrect tendencies more than the correct, and retard learning. Verbal learning frequently represents situations of this sort, with many *S–R* connections to be

[1] This article is based on a dissertation submitted to the faculty of the Department of Psychology, State University of Iowa, in partial fulfillment of the requirements for the Ph.D. degree. The writer is greatly indebted to Professor I. E. Farber for his generous guidance and criticism.

learned, and with a preponderance of wrong responses marking the early part of the learning task.

The present study is concerned with the second factor, i.e., it is an empirical investigation of the hypothesis that the effect of increased anxiety upon experimental performance is a function of the relative number and strength of incorrect response tendencies elicited in a given situation. Assuming anxiety, defined by the Taylor inventory, to have drive properties, an increase in anxiety should be accompanied by augmentation of all response tendencies. Furthermore, stronger tendencies, correct or incorrect, should show relatively greater augmentation, in strength or frequency, than weaker ones. Where incorrect tendencies are strong and numerous in comparison with correct, as in a complex verbal learning situation, anxious *S*s might show a relative decrement in performance as compared with nonanxious *S*s. In a situation where few and weak incorrect tendencies are elicited by the task, as in the usual eyelid conditioning task, anxious *S*s might show comparatively superior performance.

Assuming that the variables of intralist similarity and association value of items should affect the number and strength of correct and incorrect tendencies in a verbal learning situation, the hypothesis outlined above was tested by the use of three lists of nonsense syllables. In terms of the hypothesis, performance measures for anxious *S*s should show a relatively greater increment than those for the nonanxious *S*s as the associative value of items increased and as inter-item similarity decreased, i.e., as the ratio of potential correct to incorrect responses increased.

Procedure

Subjects.—All *S*s were undergraduates, naive concerning the experiment, and were selected on the basis of scores obtained on a modification of the Anxiety Scale devised by Taylor (**10**) as described by Rosenbaum (**8**). The Lie Scale of the Minnesota Multiphasic Personality Inventory was used to detect false low anxiety scores; a Lie score of 7 or above (the upper 10% of the Lie Scale distribution) eliminated an *S* from selection. Sixty *S*s with anxiety scores of 9 or below (lower 20%) made up the nonanxious group, while sixty *S*s with scores of 30 or above (upper 10%) made up the anxious group. The anxious group included 33 males and 27 females, the nonanxious, 45 males and 15 females.

Apparatus and method.—A Hull-type memory drum was mounted behind a gray fiberboard screen. The exposure aperture of the drum was visible through a 2 × 6-in. opening in the screen. Three lists of three-letter nonsense syllables of 12 syllables each were typed in capital letters on an endless white tape. List I used items of high similarity and low association value, List II used items of low similarity and low association value, while List III items were of low similarity and high association value. Glaze's (**3**) association values were used. The lists with their mean association values are presented below.

List I	List II	List III
FOZ	BEW	DAR
ZIS	HAJ	FUZ
MOF	SOZ	TEL
SIF	YUD	SIK
SOZ	GEX	ROP
MIF	MIP	JUN
FOS	VOH	LAF
ZOF	WUC	HIR
FIZ	XAY	MED
SOF	ZEN	BUX
ZIF	FOZ	PAC
MIZ	JIV	KOM
25.5%	42.7%	90.0%

Lists were made in four orders, all using the same sequence, but with the different orders beginning respectively with the first, fourth, seventh, and tenth syllables of the basic lists.

Syllables were presented serially. Each syllable was exposed for 2 sec. and there was an interval of 2 sec. between the disappearance of a syllable and the appearance of the next. The interval between the disappearance of the final syllable and the appearance of the first syllable for the next trial was 6 sec. Instructions were of the usual type for anticipation learning, requiring spelling of the syllables and encouraging an attempted response for each exposure.

Procedure and design.—Anxious and nonanxious *S*s were randomly assigned to each of the

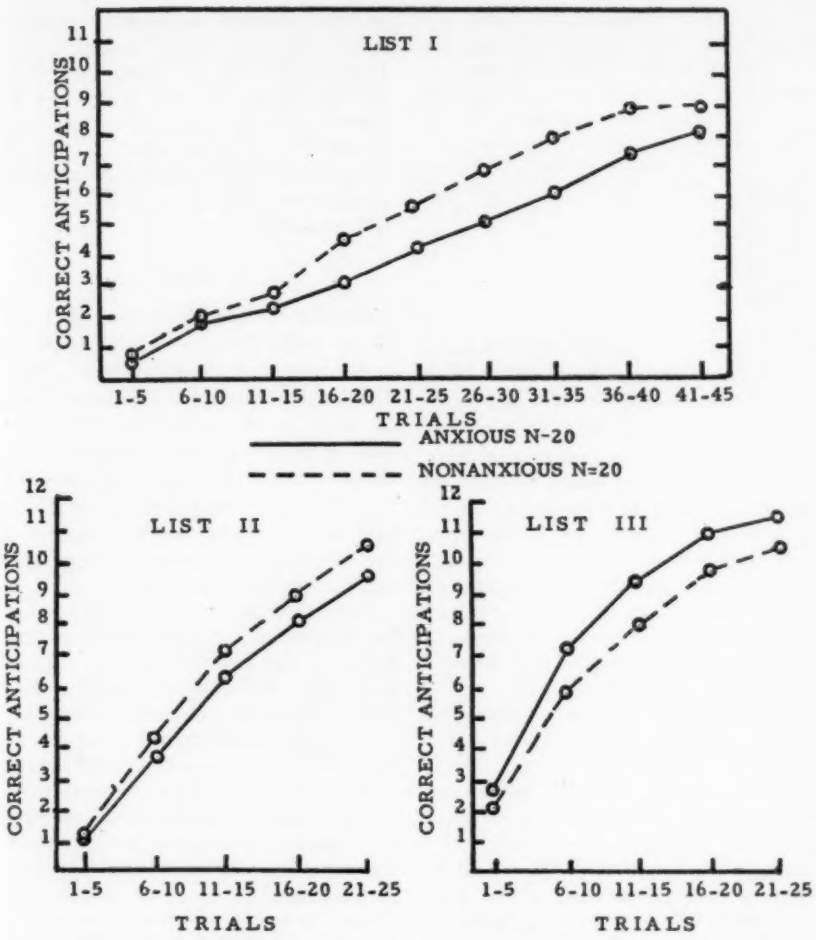
LIST I
CORRECT ANTICIPATIONS
1 2 3 4 5 6 7 8 9 10 11
1-5 6-10 11-15 16-20 21-25 26-30 31-35 36-40 41-45
TRIALS
ANXIOUS N-20
NONANXIOUS N=20
LIST II
CORRECT ANTICIPATIONS
1 2 3 4 5 6 7 8 9 10 11 12
1-5 6-10 11-15 16-20 21-25
TRIALS
LIST III
CORRECT ANTICIPATIONS
1 2 3 4 5 6 7 8 9 10 11 12
1-5 6-10 11-15 16-20 21-25
TRIALS

FIG. 1. Mean number of correct anticipations per trial over blocks of five trials

three lists, 20 of each class to each list. A simple assignment system prevented the *E* from knowing, at the time of running, whether an *S* was considered anxious or nonanxious. List I was presented for 45 trials since preliminary trials had indicated its extreme difficulty. Lists II and III were presented to the *S*s for 25 trials. In order to smooth serial position curves, *S*s in all groups were assigned in order of appearance to each of the four orders of the lists. The experimental plan presents, for analysis of variance, a two-way table with two levels of anxiety and three levels of task complexity.

RESULTS

Performance curves, plotted for anxious and nonanxious *S*s for each list, are shown in Fig. 1. Each point represents the mean number of correct anticipations per trial over a block of five trials. Comparison clearly reveals the difference in task difficulties, with List I being by far the most difficult. While Lists II and III seem

TABLE 1

MEAN NUMBER OF CORRECT RESPONSES PER TRIAL FOR BLOCKS OF 20 TRIALS

List	Trials	Anxious		Nonanxious		t
		Mean	*SD*	Mean	*SD*	
I	6–25	3.06	1.31	3.44	1.49	1.04
I	26–45	6.67	1.99	8.06	2.79	1.80
II	6–25	6.98	2.00	7.79	2.14	1.21
III	6–25	9.91	1.39	8.69	2.05	2.15*

* Significant at .05 level.

fairly similar, inspection reveals more rapid learning for List III. With List I, nonanxious *Ss* were generally superior to anxious *Ss*. This superiority is more marked after the early trials but begins to lessen as learning continues. List II also shows superiority for the nonanxious *Ss*, but to a lesser degree and with less indication of the early equality of performance that marked List I. A reversal of performance takes place in learning List III, with anxious *Ss* now becoming superior to nonanxious *Ss*.

Over-all mean scores representing the performance discussed above are presented in Table 1. The first block of five trials was excluded because of the necessary near-equality of performance for both groups during the first five trials, particularly in List I, which would obscure the actual performance difference between groups. Two sets of scores were computed for List I. The first, covering Trials 6–25, is presented in order to represent the same trial blocks from each list. The second, covering Trials 26–45, is presented to show more clearly the relations obtaining between groups after learning on List I had reached a point comparable to that seen on other lists.

The design of the study is such that a factorial analysis of variance could be used, with primary interest focused on the *interaction* between experimental lists and level of anxiety. For Trials 6–25 with all lists, the interaction F was 3.66, with 2 and 114 *df*. For the last 20 trials, the interaction F was 4.13 with the same *df*. The F ratios required for .01 and .05 levels of significance, with 2 and 100 *df*, are 4.82 and 3.09.

DISCUSSION

Results obtained indicated that the chosen tasks did differ in the variables desired and that the differences in performances of anxious and nonanxious *Ss* were in the predicted direction. The results are taken generally to support the hypothesis that anxiety has drive properties which combine with habit strength and act to increase the difference between stronger and weaker response tendencies.

However, alternative factors must be considered. Earlier mention was made of the possibility that stimuli associated with anxiety itself might elicit responses (visceral, muscular, etc.) which might interfere with learning. Thus, anxious *Ss* might react with greater feeling of frustration or failure to the complex task, thereby arousing interfering responses. While this does not explain the superior performance of anxious *Ss* on the simple task, it remains a possible factor in complex tasks. Another possible explanation is that in terms of ${}_sH_R$ rather than in terms of D alone, as suggested by Taylor (**10**, p. 89). If anxious *Ss* enter the task with greater motivation to perform well than do the nonanxious, then the making of any response, right or wrong, should result in a greater reduction of drive for the anxious *Ss*, with consequently greater increments of ${}_sH_R$ for the responses made. Then, with either correct or incorrect re-

sponse tendencies dominant, we should expect the obtained results.

Whether the results of this study are interpreted in terms of ${}_{S}H_{R}$, in terms of the augmenting role of D, or in terms of interfering responses attached to the drive stimuli of anxiety, the basic assumption presupposes drive differences between experimental groups. The results of this experiment demonstrate the interaction of drive factors with associative factors in such a way as to emphasize the importance of both in determining performance. It is felt that this study adds weight to the growing belief that both drive and associative factors must be considered in performance studies, whether in the general area of learning or in the clinical area.

Finally, the problem of determination of anxiety level is important. In the present experiment, an attempt was made to use *Ss* who obtained scores between the eightieth and ninetieth centiles on the Anxiety Scale, i.e., an intermediate range. Performances of these *Ss* were indistinguishable from those of the nonanxious groups at all points on the learning curves. The implication is plain, at least for this experiment and possibly for most healthy student *Ss*, that only the highest scores on the Anxiety Scale delineate a group different from the rest.

Summary

This study investigates the effect of anxiety upon experimental performance as a function of the relative number and strength of correct and incorrect tendencies elicited in the experimental situation.

Assuming that anxiety, as drive, combines multiplicatively with habit strength, it was hypothesized that anxiety would increase the difference between stronger and weaker response tendencies in a learning situation, aiding performance if correct tendencies were more frequent and stronger, decreasing performance if incorrect tendencies were dominant. Three verbal learning tasks, made to vary in the relative number of correct and incorrect tendencies elicited, by the manipulation of intralist similarity and association value of the nonsense syllables employed, were given to independent groups of twenty anxious and twenty nonanxious *Ss*.

Performance measures were in the predicted direction. Anxious *Ss* performed less well than nonanxious *Ss* on the difficult task with many incorrect tendencies, showed greater improvement of performance as the task became easier, and surpassed nonanxious *Ss* on the task with the least number of incorrect tendencies.

Various explanations were considered. Results were interpreted within the general framework of Hullian learning theory.

(Received July 21, 1952)

References

1. Brown, J. S., & Jacobs, A. The role of fear in the motivation and acquisition of responses. *J. exp. Psychol.*, 1949, **39**, 747–759.
2. Farber, I. E. Response fixation under anxiety and non-anxiety conditions. *J. exp. Psychol.*, 1948, **38**, 111–131.
3. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **55**, 255–267.
4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. Malmo, R. B., & Amsel, A. Anxiety-produced interference in serial rote learning with observations on rote learning after partial frontal lobectomy. *J. exp. Psychol.*, 1948, **38**, 440–454.
6. Miller, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.

7. Mowrer, O. H. Anxiety-reduction and learning. *J. exp. Psychol.*, 1940, **27**, 497–516.
8. Rosenbaum, G. Stimulus generalization as a function of clinical and experimentally induced anxiety. Unpublished doctor's dissertation, Univer. of Iowa, 1950.
9. Spence, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 690–729.
10. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
11. Welch, L., & Kubis, J. Conditioned PGR (Psychogalvanic response) in states of pathological anxiety. *J. nerv. ment. Dis.*, 1947, **105**, 372–381.
12. Welch, L., & Kubis, J. The effect of anxiety on the conditioning rate and stability of the PGR. *J. Psychol.*, 1947, **23**, 83–91.

Journal of Experimental Psychology
Vol. 45, No. 2, 1953

RESPONSE STRENGTH AS A FUNCTION OF DRIVE LEVEL AND AMOUNT OF DRIVE REDUCTION[1]

BYRON A. CAMPBELL AND DORIS KRAELING

Yale University

Miller, in several publications (**1, 2, 3, 4**), has developed a theory of learning in which drive is a strong stimulus and reinforcement is stimulus reduction. From this theory, it would appear that drive stimuli should follow the same laws as other stimuli. One of the better known characteristics of stimuli is represented in the Weber-Fechner law. This law states that equal stimulus changes are more easily detected at low levels of initial stimulus intensity than at higher levels of initial stimulus intensity. It might, therefore, be expected that a given amount of drive reduction would be more effective as a reinforcer if the initial drive were low than if the initial drive were high. In short, it might be that the Weber-Fechner relationship evidenced in psychophysical studies of stimulus discriminability may also be found in learning experiments, in which stimulus strength was used as a drive and reduction of stimulus strength used as a reinforcement.

More specifically, the primary purpose of the present experiment was to determine the reinforcing value of a constant voltage drop from different levels of initial shock. Two other purposes of the experiment were to determine the reinforcing value of different amounts of shock reduction from a constant level of initial shock, and to determine the reinforcing value of shock reduction to zero from different levels of initial shock. The details of the design are summarized in Table 1.

METHOD

Apparatus.—The apparatus consisted of a shock alley 60 in. long, 2¾ in. wide and 4 in. high. The top of the alley was of transparent Plexiglas. A strip of galvanized iron formed one side of the alley and half of its floor; another strip formed the other side and half of the floor. The floor strips were separated by ½ in. to allow faeces to fall through.

The galvanized metal plates were connected to the output of a radio transformer with a potentiometer adjustment, and were administered through a 250,000-ohm resistance in series with *S*. The *S* received the shock as long as his feet touched the two strips of galvanized iron. It was impossible for *S* to avoid shock by running along one side of the alley because of the narrowness of the alley. It was felt that this system was superior to the usual grid system (**5**) because *S* could not avoid shock by standing on grid bars with the same potential. The shocking technique seemed to have the additional advantage that *S* had less opportunity to vary the shock by changing the pressure and surface area of contact, as in the case of the grid system.

The alley was divided into three sections: a starting box 12 in. long, a middle section 36 in. long, and an end box 12 in. long. The three sections were partitioned from one another by two sliding doors. Raising the starting-box door operated a microswitch which closed the shock circuit and started a Springfield timer.

TABLE 1

EXPERIMENTAL DESIGN AND FINAL RUNNING SPEEDS

Shock in Alley	Shock in Goal Box			
	300	200	100	0
400	*1:* 19.0*	*2:* 50.4	*3:* 76.8	*4:* 79.7
300		*5:* 47.9	*6:* 63.6	*7:* 75.8
200			*8:* 74.4	*9:* 73.4

* Italics refer to group numbers, other values refer to reciprocals of time in the high-shock alley (running speeds). Each of these latter values is the mean of individual *S*s' median scores on the last five trials.

[1] The authors wish to thank Professors Fred D. Sheffield, Gregory A. Kimble, and Neal E. Miller for a critical reading of the manuscript.

The timer could be stopped by a hand-operated toggle switch. The starting box plus the middle section were called the high-shock alley, and the end box was called the low-shock goal box.

Subjects.—The *S*s were 70 male albino rats of the Sprague-Dawley strain. Their ages ranged from 120 to 150 days old. They were on ad libitum feeding, and had been tamed by being handled. Each *S* was run through the procedure in a single session.

Procedure.—The *S* was placed in the starting box of the alley, with the starting-box door closed and the goal-box door open. When the starting-box door was raised, the shock and the timer were turned on, and *S* could run to the end of the alley. When *S* crossed beneath a white line on the Plexiglas of the end box, *E* threw the toggle switch which stopped the timer, and simultaneously the door of the goal box was lowered. The *S* was left in the goal box for a period of 30 sec. If *S* did not go to the goal box within a period of 90 sec., he was removed from the alley and a new trial was begun. Each *S* received a series of 15 trials. The trials were massed, so that immediately following the removal of *S* from the goal box he was reinserted into the starting box.

Experimental design.—Nine groups of 7 animals were run in the conditions described in Table 1. A tenth group, 300–300, was run as a control. Originally an eleventh group, 100–0, was planned; however, preliminary work indicated that *S*s were not motivated to run by the low 100-v. shock. Consequently, the group was omitted.

Results

Table 1 is a summary of the data for the nine experimental groups. Latency scores (total time spent in the high-shock alley) were transformed into speed scores by multiplying the reciprocal of the time score by 100. For each *S* the median speed of the last five trials was calculated, and for each group the mean speed was determined from these median speeds. These means of median speeds for the last five trials are the data shown in Table 1. The results are discussed in accordance with the three purposes expressed in the introduction.

Acquisition curves for 8 of the 9 groups are shown. In these curves, the median speed for each trial was calculated and the trials grouped in threes to smooth the acquisition curves by plotting the median speed of each three trials. The first-trial running speeds are plotted separately in all figures to show *S*s' unconditioned running speeds to the various levels of shock. Within each group there is considerable variability in the response made to the onset of the shock stimulus. Most *S*s ran at a speed proportional to the amount of shock received, but some of them scratched and bit at the galvanized plates or ran backward, producing variability in the first-trial speed scores. Despite this variability a *t* test comparing all the first-trial speed scores of the 200–0, 200–100 groups with the 400–0, 400–100 groups gave a *p* value of .01, showing that the higher starting voltage produced significantly faster running.

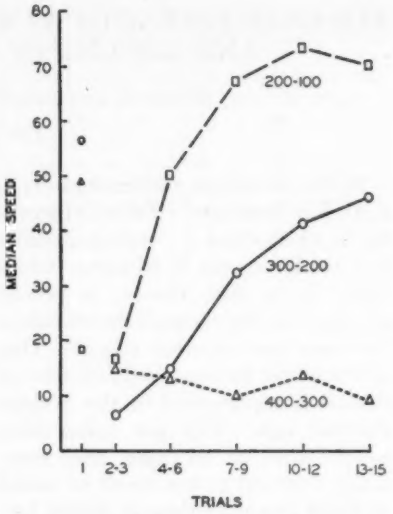

FIG. 1. Acquisition curves showing medians of median speeds plotted as a function of number of trials for groups running from various levels of shock to a shock 100 v. lower in the goal box. The first-trial running speeds are plotted separately from Trials II and III to show *S*s' unconditioned rates of running at various levels of initial shock.

(1) *Constant shock reduction—vari-*

ous levels of initial shock.—The *Ss* in these groups ran from 400 to 300 v., from 300 to 200 v. and from 200 to 100 v. In each case, there was a constant shock reduction of 100 v. The final running speeds are shown in Table 1. They show that a reduction in shock of 100 v. is more effective at a low initial shock level than at a high initial shock level. All of the differences among these groups, as computed by *t* tests, are significant ($p = .02$ level or better).

Acquisition curves are shown in Fig. 1. The curves show that the performance evidenced in the final trials is reached systematically, and that the final running speeds do not simply reflect unconditioned responses to various levels of initial shock.

(2) *Various amounts of shock reduction—constant initial shock.*—In these groups *Ss* ran from 400 to 300 v., from 400 to 200 v., from 400 to 100 v. and from 400 to 0 v. Referring to Table 1, it will be seen that *Ss* ran faster with increasing shock reduction. Differences between all groups except between the 400–100 and 400–0 groups are significant ($p = .02$ level or better).

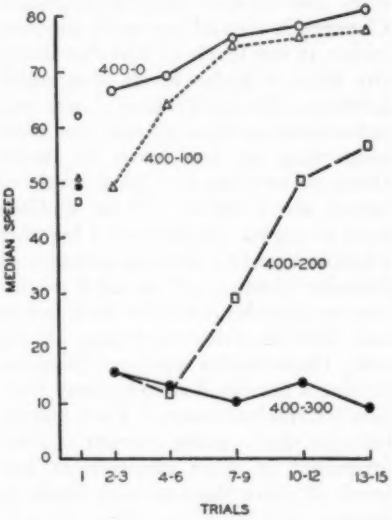

FIG. 2. Acquisition curves showing medians of median speeds plotted as a function of number of trials for groups running from the same level of shock to various levels of shock in the goal box. The first-trial running speeds are plotted separately from Trials II and III to show *Ss'* unconditioned rates of running at various levels of initial shock.

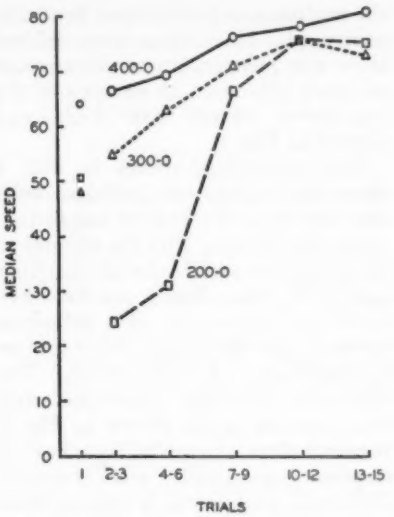

FIG. 3. Acquisition curves showing medians of median speeds plotted as a function of number of trials for groups running from various levels of shock to zero volts in the goal box. The first-trial running speeds are plotted separately from Trials II and III to show *Ss'* unconditioned rates of running at various levels of initial shock.

Again the acquisition curves are important and are shown in Fig. 2. They show that the running speeds vary systematically with the amount of drive reduction, throughout training as well as at the limit of performance.

(3) *Total shock reduction—various levels of initial shock.*—The animals in these groups ran from 400 to 0 v., from 300 to 0 v., and from 200 to 0 v. The running speed scores shown in Table 1 indicate that there are no significant differences between these groups when

the median speeds of the last five trials were used. Since these scores did not show any relationship to the amount of shock reduction, an analysis of the acquisition curves was made and plotted in Fig. 3.

The acquisition curves in Fig. 3 show that, where the goal-box voltages are zero, the rate of acquisition varies significantly with the amount of drive reduction while the final running speeds do not. When the first five trials are compared, the difference between the 400–0 and 200–0 groups is significant ($p = .01$ level). The difference between acquisition and final running speed shown in Fig. 3 suggests that an artificial ceiling of response magnitude was imposed. Following acquisition it appears that all animals run at their maximum speed. If some other measure had been taken such as resistance to extinction or strength of pull, it is possible that there would have been differences in response strength proportional to the amount of drive reduction.

If the 400–100, 300–100, and 200–100 groups are plotted and analyzed, the same general results are obtained. That is, the groups reach the same asymptote (no significant differences between groups in final running speeds) and their acquisition curves vary systematically with amount of drive reduction. In addition, t tests were calculated to compare these groups with the groups having zero volts in the end box. No significant differences in acquisition (first five trials) or final running speeds (last five trials) were found between groups having zero and 100 v. in the end box when the initial voltages were the same. This leads us to believe that the difference between zero and 100 v. in the end box is not an important stimulus change for the animal, and that it is close to the shock threshold for these animals in this apparatus.

(4) *Control.*—The tenth group running from 300 v. in the high-shock alley to 300 v. in the goal box was used as a control. These animals ran significantly slower than all other groups, except the 400–300 group, with a mean speed of 1.83 ($p = .01$ level or better), and showed a steadily decreasing speed of running.

Discussion

The most important finding is that a given amount of shock reduction is less reinforcing at high levels of initial shock than at low levels of initial shock. Thus, it can be seen that the laws of stimulus discriminability and the laws of reinforcement (drive reduction) are roughly parallel functions. Changes in stimuli are more discriminable at low levels of stimulus intensity than at higher levels of stimulus intensity (Weber-Fechner law), and reductions in drive stimuli are more reinforcing at low levels of initial shock drive than at higher levels of initial shock drive. These findings tend to negate any theory of learning which is based on absolute amounts of stimulus change, and make it necessary to consider the drive level before and after the drive reduction. However, these results were not intended to give a precise Weber-Fechner fraction for reinforcement. They merely indicate that a given amount of reinforcement is more effective at low levels of drive than at high levels of drive.

Summary

This experiment was performed to determine the reinforcing effect of given amounts of drive reduction as a function of initial level of drive. Ten groups of Ss were run from some high

initial level of shock to a lower level of shock. Three groups running from various initial levels of shock to levels of shock 100 v. lower (i.e., constant shock reduction) present the most singular important finding of this experiment. These groups show that a constant reduction in shock is more effective at a low initial shock level than at a high initial shock level. Other groups demonstrate that greater amounts of shock reduction from the same initial shock level lead to faster running speeds. A third finding is that greater amounts of shock reduction from various initial shock levels to zero volts do not influence final running speeds but do influence rate of acquisition. In general, these results confirm a theory of drive reduction, and suggest strongly that a principle like the Weber-Fechner function is applicable to the reinforcing properties of drive reduction.

(Received July 28, 1952)

References

1. Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.
2. Miller, N. E. *Notes on the concept of drive and its relationship to reinforcement.* Mimeographed Memoranda, Yale Univer., 1947.
3. Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951, 435–472.
4. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univer. Press, 1941.
5. Mowrer, O. H., & Miller, N. E. A multipurpose learning-demonstration apparatus. *J. exp. Psychol.*, 1942, **31**, 163–170.

Journal of Experimental Psychology
Vol. 45, No. 2, 1953

THE INTELLIGIBILITY OF SPEEDED SPEECH[1]

WILLIAM DAVID GARVEY

Naval Research Laboratory

From the results of previous research, it has become apparent that the speech wave pattern can be variously distorted without serious losses in the resulting intelligibility. The fact that speech can be distorted to a large degree and remain intelligible may be explained by the presence of an excess of auditory cues in the "normal" speech pattern. The listener responds most correctly when all these cues are present, but a large number of these cues can be removed without greatly affecting intelligibility. However, investigations on the effect of distortion in the form of acceleration of speech material have been meager. Fletcher (**2**) indirectly discusses the problem of acceleration in a study on the effect of frequency shifts on speech intelligibility. Frequency shifts in the speech pattern were obtained at the same time that the rate of presentation of the speech material was increased or decreased. Such increases or decreases in the rate were obtained by accelerating or decelerating speech recorded on discs or sound film. Large losses in intelligibility were not found with this method until the speed was 1.4 times that of the original speech speed.

Goldstein (**4**) reported a study in which paragraphs from several reading comprehension tests were read aloud at various rates: 100, 174, 248, 285, and 322 words per minute. High intelligibility scores (listening-comprehension scores) were obtained for the lower rates, and the *Ss* were able to understand some of the speech even at the very rapid rate of 322 words per minute.

The methods of accelerating the speech in these two studies have certain limitations. The method reported by Fletcher is limited by the distortion effects resulting from the frequency shift. Goldstein's method is limited by the rate at which the reader is able to speak, and also by the changes in enunciation and timing involved in attempting to speak rapidly.

One of the major tasks of the present project was to determine a more efficient method for accelerating speech. A possible method was suggested from the general findings of studies on the effects of various distortions on speech intelligibility. Particularly from Miller and Licklider's work (**6**) on interrupted speech, it appears that the undistorted speech pattern affords an excess of cues, which may be utilized by the listener but are not essential to intelligibility. In their work the speech pattern was interrupted systematically during the course of its presentation to the *Ss*. Such an interruption was accomplished by means of an electronic switch which turned the speech on and off at a desired rate. These authors reported that when the interruptions

[1] This article is based on a thesis presented in partial fulfillment of the requirements for the M.A. degree, Psychology Department, University of Virginia, 1949. The author wishes to acknowledge the assistance of Dr. W. R. Thurlow, who directed this research, and of Dr. R. H. Henneman, who directed an Air Force contract of which this research was a project. This contract, No. W33–038–ac–21269, was sponsored by Aeromedical Laboratory, Wright Air Development Center.

The opinions or assertions contained here are the private ones of the writer and are not to be construed as official or reflecting the views of the Air Force.

were made at a rate of ten times per second, intelligibility of monosyllabic words did not drop below 90% until over 50% of the original speech pattern had been removed.

In Miller's and Licklider's electronic chopping technique, the gaps in the speech record were not closed after chopping; hence no acceleration was achieved. It appeared possible, however, that if some means could be devised for removing the silent spaces, thus leaving an abbreviated speech record, an increase in the speech rate with no great losses in intelligibility for small "chops" might result. This method, moreover, would not have the disadvantage of shifting the frequency of the speech spectrum with acceleration.

Procedure

Method of accelerating speech material.—As a source of words for the experiment the Harvard PAL Auditory Test No. 14 (composed of spondaic words such as *baseball* and *sunset*) was used. The words from this test were recorded on a plastic base tape with a magnetic recorder. To replace the standard pulling mechanism on the recorder, a larger capstan was made by means of which approximately 40 cm. (instead of the standard 19.5 cm.) of tape per second passed the magnetic recording head. This alteration enabled *E* to have greater lengths of tape with which to work in the chopping technique. The position of the words on the tape was determined by slowly passing the tape over the magnetic playback head of the recorder. The beginning and ending of each word could be determined auditorily and then marked on the tape. Once the position of the words had been determined, the sections to be removed were marked on the tape and later cut out. The ends of the segments remaining to form a new abbreviated speech pattern were then spliced together. An acceleration of twice the original speed could be accomplished by removing one centimeter of the record, leaving in the second centimeter, removing the third centimeter, leaving in the fourth, and so on throughout the length of the speech record.

In order to be systematic in chopping the test words, the first chop to be removed was always the first segment at the beginning of the speech record. For example, an acceleration of 4.0 times the original speed was achieved by removing the first three centimeters of the speech pattern on the record and leaving the fourth centimeter of the pattern to form the first centimeter of the new abbreviated word. This procedure was repeated throughout the length of the recorded word.

After the tapes had been chopped at the various rates, the abbreviated records were transcribed on a second intact tape. The recording intensity was maintained at a constant level by controlling the input-output ratio of the volume controls and the primary voltage on the line. These new intact abbreviated tapes were used to present the test materials to *S*. The test material was presented monaurally over Permoflux PDR No. 10 earphones in a partially soundproof room at an intensity 44 db above a level at which *S*s could just detect the speech sound. Preliminary experimentation indicated that this was an intensity at which *S*s could easily understand unaccelerated words, and since subsequent experiments were to be conducted to determine the effect of various intensities on the intelligibility of accelerated speech this sound level was used. To the *S*s the test words merely appeared to have been accelerated without apparent frequency translation and without distracting clicks or other indications of gaps between splicing.

Design.—In order to examine as fully as possible the variables involved in this "chop-splice" accelerating technique it was decided to investigate the relationships between intelligibility and (*a*) the general effect of acceleration, (*b*) the size of the chop removed when acceleration was held constant at twice the original speed, and (*c*) the percentage of the speech pattern removed.

To investigate these relationships, 11 experimental tapes were used, each with a different relationship between the removed and remaining segments. The sizes of the removed and remaining segments with attendant accelerations are shown in Table 1. "Chop-splice" segments are expressed in length of tape, which may be interpreted in terms of time, since 1 sec. is represented by .025 cm. of tape. The relationship between acceleration and intelligibility was determined by using all 11 experimental record tapes. Thus a range of accelerations from 1.5 to 4.0 times the original speed was obtained.

In ascertaining the relationship between intelligibility and the size of the chop removed, with acceleration maintained at a constant 2.0 times the original speed, the data resulting from Tapes 2, 4, 7, and 10 were used.

The results obtained from investigating these above variables differed greatly from previous investigations of speech accelerated with an accompanying frequency shift. It, therefore,

TABLE 1

DESCRIPTION OF TAPES AND RESULTING MEANS AND STANDARD DEVIATIONS OF INTELLIGIBILITY SCORES FROM ACCELERATION WITHOUT FREQUENCY SHIFT ($N=6$ CASES PER TAPE)

Tape	Size of Segment		Amount of Acceleration	% of Word Removed	Intelligibility	
	Removed (cm.)	Remaining (cm.)			Mean	*SD**
1	1.0	2.0	1.5	33.3	98.67	1.49
2	1.0	1.0	2.0	50.0	95.33	2.23
3	1.5	2.0	1.8	42.9	95.67	3.90
4	1.5	1.5	2.0	50.0	95.67	2.34
5	1.5	1.0	2.5	60.0	93.33	3.39
6	2.0	1.0	3.0	66.7	78.33	13.05
7	2.0	2.0	2.0	50.0	95.00	2.52
8	2.0	3.0	1.7	40.0	93.33	2.26
9	2.5	1.0	3.5	71.4	58.00	6.38
10	2.5	2.5	2.0	50.0	85.67	3.91
11	3.0	1.0	4.0	75.0	40.00	6.00

* $SD = \sqrt{\frac{\Sigma x^2}{N}}$

appeared advisable to investigate the effect of acceleration (with an accompanying frequency shift) on intelligibility of spondaic words. Five additional experimental tapes were prepared (see Table 2), using the same 50 spondaic test words, and were presented under the same general experimental conditions. In place of chopping and splicing the words to obtain the desired accelerations of 1.5, 1.7, 1.8, 2.0, and 2.5, the entire list of words was presented to *Ss* at the desired rates by merely increasing the speed of the speech record across the recording pickup. For this method of acceleration, it was necessary to prepare tapes with additional blank intervals spliced between the words in order that the rate of word-stimulus presentation might be held constant. With other experimental conditions held constant, it was thus possible to make a direct comparison between intelligibility scores resulting from conditions of accelerated speech with the frequency shift and scores resulting from conditions of accelerated speech by the "chop-splice" technique without an accompanying frequency shift.

General experimental situation.—Six *Ss* were

TABLE 2

DESCRIPTION OF TAPES AND RESULTING MEANS AND STANDARD DEVIATIONS OF INTELLIGIBILITY SCORES FROM ACCELERATION WITH FREQUENCY SHIFT ($N=6$ CASES PER TAPE)

Tape	Amount of Acceleration	Accompanying Frequency Shift*	Intelligibility	
			Mean	*SD*†
12	1.5	1.5	97.0	2.24
13	1.7	1.7	90.0	5.00
14	1.8	1.8	88.3	3.90
15	2.0	2.0	65.0	7.63
16	2.5	2.5	8.0	11.40

* Multiple by which over-all frequency spectrum of the speech pattern is increased.

† $SD = \sqrt{\frac{\Sigma x^2}{N}}$

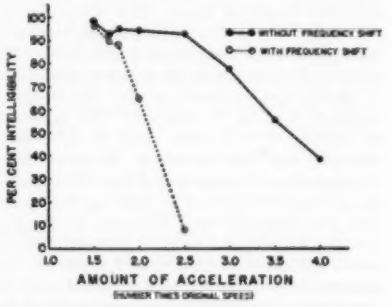

FIG. 1. The relationship between intelligibility and the amount of acceleration. The solid curve represents results obtained from accelerating speech without an accompanying frequency shift. The dotted curve represents the results obtained from accelerating speech with an accompanying frequency shift.

assigned randomly to each of the experimental conditions described in Tables 1 and 2. The *Ss* were 96 male college students in either an introductory or an experimental psychology course at the University of Virginia. In order to insure that the intelligibility scores were not influenced by faulty hearing, all *Ss* were tested before the experiment proper for auditory acuity with the Harvard PAL Auditory Test No. 12. The acuity test was administered under the conditions prescribed by Hudgins *et al.* (5).

To insure uniformity, all instructions were recorded and played to *Ss*. At the end of the instructions, *S* was given an opportunity to ask questions. After this period a warning was given and the accelerated speech material was presented. After each presentation of a word, a standard time was given *S* to respond with the word that had been heard. These responses were recorded on discs and later scored for accuracy by *E*.

Results

The general effect of "chop-splice" acceleration on speech intelligibility.—The *Ss* showed agreement within any single condition of acceleration, yet there were some significant differences among experimental conditions. The means and *SD*'s of the scores obtained from each experimental condition are presented in Table 1.

The graph in Fig. 1 shows the relationship between intelligibility and the amount of acceleration. The solid curve is a plot of the data obtained using the "chop-splice" technique. It can be seen that 95% or greater intelligibility was obtained for accelerations which were 1.5, 1.75, and 2.0 times the original speed of the speech material. For speeds of 1.67 and 2.5, mean intelligibility scores above 90% were obtained.

At an acceleration of 3.0 times the original speed, a mean intelligibility score of 78% was still obtained. It was not until the test words had been accelerated to 3.5 times the original speed that the resulting mean intelligibility score approached the 50% level. Moreover, the mean intelligibility score did not drop below the 50% level until an acceleration to 4.0 was attained.

It would appear that with the acceleration technique used in this experiment, using discrete spondaic test words, the critical speed is around 2.5 times the original speed, for it is not until this point is exceeded that the intelligibility scores begin to show a significant drop.

The intelligibility scores as a function of amount of acceleration were tested for statistical significance with Friedman's (3) ranking test. The conditions which were to be tested were arranged in two-way columns, and the scores in each column were arranged in accordance with the respective scores each *S* received on the auditory acuity test. The rows (across columns) were next ranked in accordance with the scores received on the experimental conditions. The level of significance is the per cent of time one would expect to get by chance a chi-square value (χ^2) equal to or greater than the computed χ_r^2. The test revealed that the scores were significant beyond the 1% level of confidence.

Relationship between intelligibility and size of chop removed (with amount of acceleration held constant).—*t* tests were made on the differences between intelligibility scores resulting from Tapes 2, 4, 7, and 10, and it was found that the scores from Tapes 2, 4, and 7 did not differ significantly from each other. However, the scores resulting from Tape 10 did differ significantly from the scores of Tapes 2, 4, and 7 at the 1% level of confidence. This significance level was tested with Festinger's (1) ranking test and found to be the same. These data indicate that when the acceleration is held constant at twice the original speed, the size of the chop removed does not

significantly affect intelligibility until it is at least 2.5 cm. in length.

Relationship between intelligibility and percentage of speech pattern removed.—Although an increase in percentage of speech pattern removed is accompanied by a proportional increase in acceleration, it is of interest that such large percentages of the speech pattern can be removed, even with an accompanying increase in speed, before intelligibility is greatly reduced. The mean intelligibility scores do not drop below 80% until 67% of the speech pattern is removed. It is not until 75% of the speech pattern is removed that scores drop below 50%.

Comparison of intelligibility scores obtained from accelerating speech with and without frequency shifts.—The resulting mean intelligibility scores and standard deviations obtained from accelerating discrete spondaic words with an accompanying frequency shift are presented in Table 2. A comparison of the two methods of acceleration can be made from the graphs in Fig. 1.

It is clear from this figure that the attendant frequency shift plays an important role in reducing the intelligibility of speeded speech. At an acceleration of 1.75, using the same test words and identical experimental procedure, the mean of the intelligibility scores, with the concomitant frequency shift, was 90%; for the same degree of acceleration without frequency shift the mean intelligibility score was 95%. For an acceleration of 2.0, mean intelligibility with the frequency shift was 65%, but was 95% with the "chop-splice" technique. When the acceleration was increased to 2.5, a mean intelligibility score of less than 10% was obtained with the frequency shift, and a mean intelligibility score of 93% was obtained with the "chop-splice" technique.

t tests were made of the differences between the mean intelligibility scores of the two methods at the various acceleration levels. The differences between the scores obtained for accelerations of 1.5 and 1.67 for the two methods were not found to be statistically significant. For the acceleration of 1.75 the difference was found to be statistically significant at the 5% level; for speeds of 2.0 and 2.5 the differences were found to be significant beyond the 1% level. These levels of significance were checked with Festinger's (1) ranking test and the same levels of significance were found.

Discussion

In accelerating speech by the "chop-splice" technique the number of speech cues is greatly reduced; yet speech remains intelligible until well over 50% of the record is removed. These results further support the hypothesis drawn from the earlier experiments on distortion, that speech contains more cues than are necessary for good intelligibility under normal listening conditions.

When speech is accelerated by the "chop-splice" technique, intelligibility is probably a function of both acceleration and the amount of the speech record removed. There is a problem, however, in determining whether decreases in intelligibility are caused primarily by an acceleration of the remaining sounds or by the removal of sounds. To answer this question, we need to know the extent of intelligibility lost by merely removing parts of the speech pattern in a word without an attendant acceleration.

Miller and Licklider (6), in using their electronic chopping technique, removed certain amounts of the speech pattern without an accompanying acceleration. A cursory experiment

was performed by the author using the "chop-splice" technique on 20 of the same monosyllabic words which Miller and Licklider used. With the exception of the acceleration which accompanied the removal of portions of the speech material, the experimental conditions were maintained as nearly as possible identical with those used by Miller and Licklider. The 20 test words were chosen randomly from Miller and Licklider's list. Two experimental record tapes were "chopped and spliced," one which had 62% of its speech material removed (acceleration of 2.7) and another which had 50% of its speech material removed (acceleration of 2.0).

The results of the experiment indicate that when 50% of the speech pattern is removed, there is no appreciable difference in the mean intelligibility scores for the two methods, i.e., with and without an accompanying acceleration. The fact that there is no difference at this point may be attributable to a distortion which accompanied Miller and Licklider's electronic chopping. This distortion which was caused by the silent spaces between the speech record, although somewhat present, was not reportedly distracting in the "chop-splice" technique described in this article. Miller and Licklider state that, owing to this distortion, the talker sounded as though he were speaking in a "hoarse and raucous" voice. Such apparent distortion may have balanced out any advantage which the nonspeeded chopped speech may have had over the speeded chopped speech. Also it is possible that a speed of 2.0 times that of normal speech rate may not be a rate which is sufficient to cause a detrimental effect on intelligibility. When 62% of the speech pattern is removed, there is a striking difference (a difference of 40%) between the mean intelligibility scores of the accelerated and nonaccelerated speech in favor of the nonaccelerated speech. Assuming that the remaining experimental conditions were equivalent for the two experiments, this difference must be attributable to the difference in presentation rate. It is, therefore, reasonable to conclude tentatively that the rate of the speech sounds (at least for the more rapid rate of 2.7) as well as the amount of the speech pattern removed is an important variable affecting the intelligibility of "chop-splice" speeded speech.

Summary

1. A new technique for accelerating speech was developed, based on chopping out portions of the speech record on a plastic base recording tape, then splicing together the remaining segments of the record tape to form a new intact record.

2. With this "chop-splice" accelerating technique, using discrete spondaic words, accelerations as great as 2.5 times the original speech speed were attained with less than 10% loss in intelligibility.

3. When acceleration was held constant at twice the original speed, the size of removed chops did not have a significant effect upon intelligibility until 2.5 cm. chops were employed.

4. A comparison of the "chop-splice" technique with an accelerating technique, which had an accompanying frequency translation, indicated that with accelerations greater than 1.67 times the original speech speed the distortion resulting from the frequency translation had a greater effect upon intelligibility than did the distortion resulting from the removal of portions of the speech pattern with the "chop-splice" accelerating technique.

(Received July 24, 1952)

References

1. Festinger, L. The significance between means without reference to the frequency distribution function. *Psychometrika*, 1946, **11**, 97–106.
2. Fletcher, H. *Speech and hearing*. New York: Van Nostrand, 1929.
3. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. statis. Ass.*, 1937, **32**, 675–701.
4. Goldstein, H. Reading and listening comprehension at various controlled rates. New York: Teachers College, Columbia Univer., 1940.
5. Hudgins, C. V., Hawkins, J. E., & Karlin, J. E. The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope*, 1947, **57**, 57–89.
6. Miller, G. A., & Licklider, J. C. R. The intelligibility of interrupted speech. *J. acoust. Soc. Amer.*, 1950, **22**, 167–173.

Journal of Experimental Psychology
Vol. 45, No. 2, 1953

INTRALIST–INTERLIST RELATIONS IN VERBAL LEARNING

DONALD R. MEYER AND R. C. MILES

The Ohio State University

Analysis of successive transfer situations has been one of the least popular endeavors in psychology of learning. Yet recent attempts to deal with data of this kind have pointed to the great significance of the topic as a source of materials for theoretical bridges to psychology of thinking. The theories are particularly significant in that they permit continuous interpretations of discontinuous and apparently discontinuous acquisition, and thus provide a resolution for a classical controversy.

One interpretation of changes that take place when an individual practices with several examples from a problem class has been advanced by Bateson (1). He postulates that there are two gradients in learning. The first is represented by rote-learning curves and has been labeled proto-learning. The second is a progressive change in the rate of proto-learning, and this Bateson calls deutero-learning. Although the hypothesis is highly specific, the data required for a direct test have never been collected.

There are, in fact, very few human-learning experiments that are even remotely related to the question. Perhaps the most notable example is the study of Ward (9), who showed that trials to reach criterion performance on a given list of nonsense syllables decrease as a negatively accelerated function of practice with 16 lists. Melton and von Lackum (5) found that such interlist effects could still be demonstrated after subjects had worked for 28 days with lists of syllables, words, and numbers. Similar changes have also been noted by Thune (8), who studied warm-up and learning-to-learn phenomena during presentations of three nonsense-syllable lists per day for a period of five days.

Most of the successive transfer data obtained in these experiments are extra dividends, for the procedures were designed primarily to answer other questions. As a result, the data have been presented in forms which preclude determinations of the way in which intralist learning curves change so that it is possible for there to be an interlist function.

At present, the only data susceptible to such analysis come from studies of learning by monkeys, and particularly the experiments of Harlow (3) which led to the development of the theory of learning sets. When compared to the massive amount of information on interproblem transfer now available from lower primate research, the human material is almost incredibly scanty. This experiment has been designed as a partial remedy.

Conceptually, it is very simple: a large number of people practice a nonsense-syllable list for a fixed number of trials, and then go on to another similar list for a similar number of trials. This is repeated until 20 lists have been learned. There are several ways in which the data can be ordered. Performance per list as a function of lists is one, and past experiments show that such scores get better and better. Performance on a given list as a function of practice with that particular list can also be studied. If both

approaches are used, then changes in intralist learning can be related to the amount of interlist practice.

A performance has been chosen which yields intralist learning curves that can be described initially by a positive growth function. Under these circumstances, several alternative interlist modifications are possible. One, as Bateson suggests, is that the limits of the curve remain fixed and that the rate alone varies. Another, suggested by data from monkeys (6), is that transfer appears as a discontinuous increment in performance which affects the intercept and nothing else. A final simple alternative is that the upper limit changes. Of course, a combination of changes may appear. If these possibilities fail to account for the results, the function itself must be discarded.

Method

Subjects.—The 64 *S*s were members of two sections of an elementary course in psychology. List-learning sessions took place in the final 5 min. of each daily class period. One section of 30 *S*s met at 8:00 A.M., and one of 34 met at 9:00 A.M.

Materials.—A series of 20 nonsense-syllable lists, each 12 units long, was selected from a group designed by Melton (4) to yield a minimum of interlist interference. The average Glaze association value was about 27. The lists were typed in pica capitals, double spaced, and reproduced by mimeograph on 5 × 8-in. sheets of paper. These sheets and stiff white cards of about the same size were made up into booklets. Twenty separate booklets were prepared for each *S*, one booklet for each list. Booklets were presented in 20 different random orders so that the over-all effects of possible differences in difficulty were minimized.

Procedure.—Booklets were distributed face down before each class period began. At the initiation of the list-learning session, each *S* turned over his booklet and studied the list that appeared on the first page. The command "recall" was given after 30 sec., and each *S* turned to the first blank card. A period for reproduction of syllables followed which was also 30 sec. long. Then, at the command "turn" each *S* folded over the recall card and studied the list again, which was reproduced on the third page of the booklet. This procedure was repeated five times. The *S*s were informed that it was not necessary for them to reproduce the syllables in order, but only those syllables which were wholly correct were scored. There was some variation in the time intervals between lists due to occasional absences and week-end holidays.

Results[1]

Figure 1 shows the over-all performance of the group on each of the 20 successive lists. A point on this graph represents the mean number of syllables recalled in all five trials. This score was about 24 on List 1, but by List 20 had grown to almost 35.

[1] The statistical analysis was devised by D. R. Whitney, and performed by the Statistics Laboratory of the Ohio State University. The authors wish to acknowledge this help with thanks.

TABLE 1

Means and Variances for Performances on Each Test List Trial

List	Trial									
	1		2		3		4		5	
	M	s^2	*M*	s^2	*M*	s^2	*M*	s^2	*M*	s^2
1	2.91	1.87	4.13	2.89	5.08	3.60	5.63	4.58	6.28	5.73
7	3.67	1.72	5.55	2.28	6.41	3.93	7.09	3.55	7.95	3.89
13	3.98	2.27	5.88	2.80	7.02	2.80	7.42	4.21	8.28	4.61
19	4.13	1.54	6.11	3.44	7.11	3.29	7.88	2.98	8.73	3.79

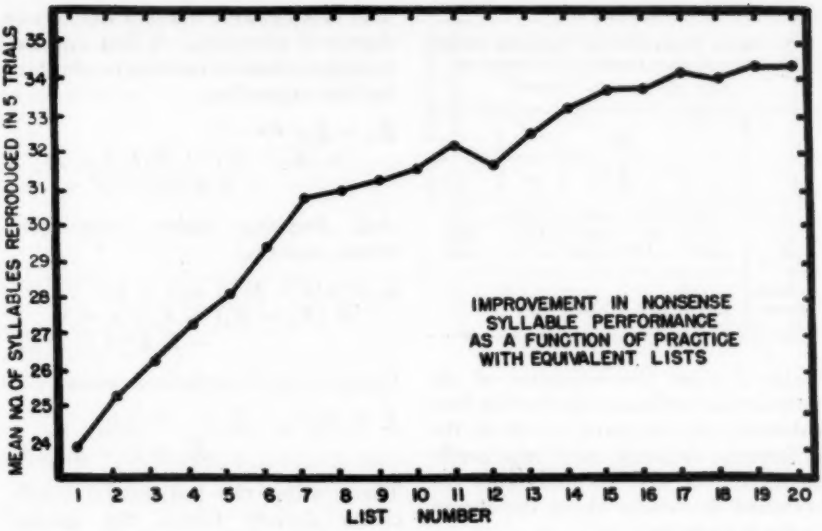

FIG. 1. Improvement in nonsense-syllable performance as a function of practice with equivalent lists. Each point represents the group mean number of syllables reproduced on all five trials with each list.

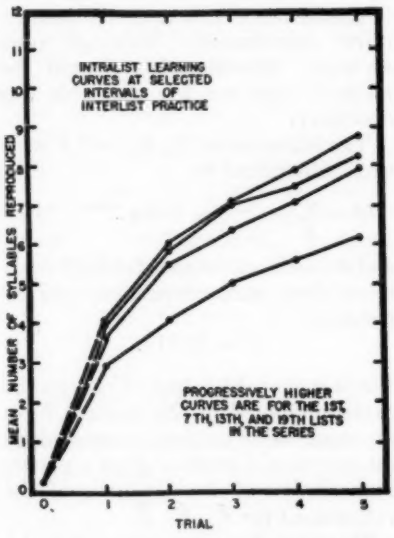
INTRALIST LEARNING CURVES AT SELECTED INTERVALS OF INTERLIST PRACTICE
MEAN NUMBER OF SYLLABLES REPRODUCED
PROGRESSIVELY HIGHER CURVES ARE FOR THE 1ST, 7TH, 13TH, AND 19TH LISTS IN THE SERIES
0 1 2 3 4 5
TRIAL

FIG. 2. Intralist learning curves at selected intervals of interlist practice. Five practice lists intervened between each of the lists depicted.

As past experiments would lead us to predict, the curve is negatively accelerated.

A family of intralist learning curves is shown in Fig. 2. From the 20 possible lists four have been selected; the first, seventh, thirteenth, and nineteenth. These test lists are separated by blocks of five lists that are ignored in this analysis. Table 1 gives the exact means and variances for the test list trial scores. It is with these data that we are primarily concerned.

First, a curve was fitted to the five empirical measures of performance on each test list by the method of orthogonal polynomials (**2**). The degree of the curve was extended until the coefficient was not significantly different from zero. In each instance the function was of the second degree,

$$y = a_0 + a_1(x - 3) + a_2(x - 3)^2.$$

TABLE 2

POLYNOMIAL COEFFICIENTS, STANDARD ERRORS OF DIFFERENCES BETWEEN COEFFICIENTS AND *SD*'S ABOUT CURVES

	List				$s_{\hat{a}_\alpha - \hat{a}_\beta}$
	1	7	13	19	
$\hat{a}_0$	5.02	6.45	6.92	7.13	.712
$\hat{a}_1$	.825	1.01	1.01	1.06	.102
$\hat{a}_2$	−.109	−.157	−.200	−.151	.086
s about curve	1.94	1.77	1.84	1.75	

Table 2 gives the estimates of the polynomial coefficients in the first four columns, the standard errors of the differences between any two coefficients in the last column, and the standard deviations about the curves in the bottom row.

Comparisons of the various coefficients indicate that those of the first and second order do not differ from curve to curve. However, the zero-order coefficient for List 1 is significantly less than the corresponding coefficients in the other instances, and there is a trend of increasing values. An implication of these findings is that the slopes of the curves do not differ, for the relevant expression is

$$y' = a_1 + 2a_2(x - 3).$$

The polynomial coefficients were then employed to estimate the values of constants for growth functions fitted to the same four sets of data, equations of the form

$$y = K_0 - K_1 e^{-Kx}.$$

For convenience in computation, the equations were written with exponents of $(x - 3)$ rather than x.

The problem of finding maximum likelihood estimates of K_0, K_1, K is equivalent to solving three transcendental equations in these unknowns. This can be done numerically to any degree of accuracy. A first approximation to these values can be obtained by first expanding

$$K_0 - K_1 e^{-K(x-3)} = (K_0 - K_1) + K_1 K(x - 3) - K_1 K^2 (x - 3)^2 + \ldots$$

and, dropping higher insignificant terms, equating

$$a_0 + a_1(x - 3) + a_2(x - 3)^2 \doteq (K_0 - K_1) + K_1 K(x - 3) - K_1 K^2 (x - 3)^2/2.$$

Comparison of coefficients gives

$$\hat{K} \doteq 2\hat{a}_2/\hat{a}_1, \quad \hat{K}_1 \doteq -\hat{a}_1^2/2\hat{a}_2, \quad \hat{K}_0 \doteq \hat{a}_0 - \hat{a}_1^2/2\hat{a}_2.$$

Then, using the polynomial coefficients already found, the needed approximations are obtained. That these values reasonably satisfy the set of maximum likelihood equations (which are also least square equations in this instance) can be determined by direct substitution. Although more accurate determinations could be made, it was not felt that this was necessary.

The estimates of K_0, K_1, and K were then substituted in

$$Q = \sum_{ij} (y_{ij} - K_0 + K_1 e^{-K(x-3)})^2$$

and an estimate of the standard deviation about each curve computed by means of

$$s^2 = Q/317.$$

The fact that the values of s about the curve computed in this manner have the same value as those computed by the polynomial method gives a partial confirmation of the adequacy of the values used for $\hat{K}_0$, $\hat{K}_1$, $\hat{K}$.

The procedure for estimating the standard error of the estimate of K was derived from the fact that the joint distribution of four maximum

likelihood estimates, $\hat{K}_0$, $\hat{K}_1$, $\hat{K}$, $\hat{s}^2$ is a 4-variate normal distribution for large n. The variance of K is the appropriate element in the inverse of the matrix

$$s^{ij} = -E\left[\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\log f(x;\theta_1,\theta_2,\theta_3,\theta_4)\right]$$

where f is a normal density with mean $K_0 - K_1\epsilon^{-Kx}$ and variance s^2, θ_i being the ith unknown parameter.

The expressions for the four curves, estimates of their standard deviations, estimates of the standard deviations of the sample estimates of K, estimates of the standard error of the difference between K's, and the values of y at $x = 0$ are summarized in Table 3. It is apparent that we cannot reject the hypothesis that the K's for all curves are the same. Furthermore, the value of the extrapolation to zero does not cluster, but instead becomes progressively larger with practice. Such changes, however, are small.

The only simple assumption which remains is that the principal change is in the upper limit or asymptote. In Fig. 3 the two extreme intralist curves, those for List 1 and List 20, are plotted in semilogarithmic coordinates with different rounded asymptotes derived from the empirical equations. With upper limits of 8 and 11

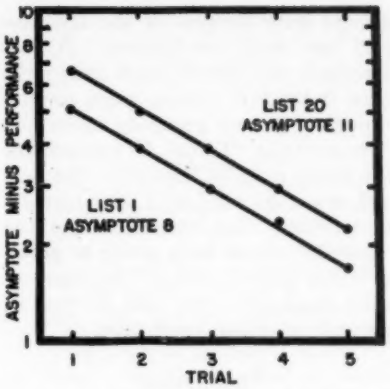

FIG. 3. Semilogarithmic plots of intralist learning curves. Assumed asymptotes are 8 and 11 for Lists 1 and 20 respectively.

for the first and last lists respectively, the plots are both rectilinear and parallel. Note that in each instance the asymptote is less than the number of syllables in the list.

Statements thus far must be limited to group learning curves. Sidman (7) has pointed out that the functions derived from pooled individual positive growth functions will be of this same form only if rates are constant from S to S. Although pooling cannot be dispensed with entirely in the treatment of data of this sort, comparisons were made to see if rate and asymptote were related. The overall score of each S was computed, and

TABLE 3

SUMMARY OF FITTED POSITIVE GROWTH FUNCTIONS

List	Equation	s About Curve	s_K	y at $x=0$
1	$y=8.13-3.11\epsilon^{-.365(x-3)}$	1.94	.157	0.51
7	$y=9.70-3.25\epsilon^{-.311(x-3)}$	1.77	.116	1.44
13	$y=9.49-2.57\epsilon^{-.394(x-3)}$	1.84	.121	1.57
19	$y=10.87-3.74\epsilon^{-.284(x-3)}$	1.75	.110	2.13
	$s_{\hat{K}_\alpha-\hat{K}_\beta} = .180$			

TABLE 4

MEAN INTRALIST PERFORMANCES FOR BEST AND WORST TEN SUBJECTS

Trial	List 1		List 20	
	Best Ss	Worst Ss	Best Ss	Worst Ss
1	4.1	1.5	5.9	3.3
2	6.5	2.2	8.1	4.5
3	7.9	2.8	9.6	4.8
4	8.5	3.4	11.6	5.7
5	9.1	3.1	11.4	6.5

groups were formed of the ten best and ten worst performers. It was assumed that within each group the true individual curves were similar enough to permit grouping under any circumstance. Table 4 summarizes the mean intralist scores for the best and worst *Ss* on List 1 and List 20. It is evident that there is no marked variation in rate from group to group at either interlist point. Asymptotes vary markedly, but this is not an obstacle to group treatment.

Discussion

It is evident that the obtained data do not support Bateson's hypothesis. In spite of the fact that there is a marked change in performance with interlist practice the rate constants do not vary significantly from one intralist function to another. Nor is there evidence for a trend in this direction.

The lump transfer hypothesis fares little better. Certainly the obtained curves are not markedly discontinuous, so if there is such an effect it is not a major factor. Actually the data suggest that there is a minor and progressive discontinuity, and so it does not follow that a curve fitted to the empirical points must include the zero point as well.

The tentative conclusion that the principal change is in the asymptote seems ludicrous at first glance. If it is true that early in interlist practice the limit is less than the number of syllables, this hypothesis implies that subjects could not learn the list in a finite number of trials. Obviously this cannot be so. However, it is possible that each practice trial changes both the amount that is learned and the amount that can be learned. From this it follows that the present empirical equations, despite the accuracy of description, are valid approximations only if it is assumed that within five trials the changes in the asymptote are negligible. This could not hold were the intralist practice period more extensive, and hence it must be concluded that a set of points obtained when learning is to criterion actually represents an envelope instead of a single curve.

Perhaps the results obtained are peculiar to the method of whole presentation. It may be that as learning proceeds the *Ss* are actually exposed to larger and larger groups of syllables. If, then, it is postulated that the elements in each group are learned at a rate proportional to the number which remain to be learned, the present data demonstrate that this rate is independent of the over-all number of elements in the momentarily effective group.

Summary

This experiment was designed to ascertain the changes that take place in the learning of lists of nonsense syllables when practice with a series of 20 lists is given. A group learning situation was employed, and the material was presented to 64 *Ss* in booklet form. The method was that of whole presentation and reproduction.

A statistical analysis of intralist learning curves for four representative lists revealed that the growth constants for fitted positive growth functions do not vary significantly with interlist practice, and that there is but slight evidence for the development of discontinuities. The data thus do not support Bateson's hypothesis that deutero-learning is a change in the rate of proto-learning, and indicate that lump transfer is a factor of little or no importance in this situation.

Intralist learning curves at different stages of interlist practice differ principally in asymptote. Since the total

amount which can be learned is a function of the amount that has already been learned, there are serious reservations upon the applicability of monotonic functions to data of this variety.

(Received July 28, 1952)

References

1. Bateson, G. Social planning and the concept of "deutero-learning." In T. M. Newcomb and E. L. Hartley (Eds.), *Readings in social psychology*. New York: Holt, 1947.
2. Fisher, R. A., & Yates, F. *Statistical tables for biological, agricultural, and medical research*. London: Oliver and Boyd, 1948.
3. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
4. Melton, A. W. Unpublished manuscript. Summarized by Hilgard, E. R., Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
5. Melton, A. W., & Von Lackum, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
6. Meyer, D. R. Food deprivation and discrimination reversal learning by monkeys. *J. exp. Psychol.*, 1951, **41**, 10–16.
7. Sidman, M. A note on functional relations obtained from group data. *Psychol. Bull.*, 1952, **49**, 263–269.
8. Thune, L. E. Warm-up effect as a function of level of practice in verbal learning. *J. exp. Psychol.*, 1951, **42**, 250–256.
9. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, Whole No. 220.

CONDITIONING AND EXTINCTION AS A FUNCTION OF ANXIETY[1]

KENNETH W. SPENCE AND I. E. FARBER

State University of Iowa

In two previous experiments (**6, 7**) conducted in the Iowa laboratory, it was found that the level of performance in eyelid conditioning as measured by the frequency of CR's during the training period was definitely related to the scores made by the *S*s on a manifest anxiety scale. Similar results have been obtained by other investigators (**1, 4, 8, 9**). A recent study by Hilgard, Jones, and Kaplan (**2**), however, failed to confirm this finding so far as the conditioning data were concerned. In a second part of their study involving differential conditioning a significant difference was obtained, *S*s with high anxiety scores giving significantly more CR's to the negative stimulus than those with low anxiety scores.

The present experiment presents further data comparing the conditioning performances of anxious and nonanxious *S*s. In addition to the frequency measure of conditioning, a measure of resistance to extinction was also available. After 60 conditioning trials the time interval between the CS and UCS was increased from 500 msec. to 2500 msec. McAllister (**3**) has recently demonstrated that the CR extinguishes (decreases) under this procedure, requiring just about as many trials to reach the level of response prior to conditioning as the number of trials involved in the original conditioning. This technique has an advantage over the usual extinction procedure in defense conditioning in that it maintains the drive level of *S*. That is, *S* continues to receive the noxious UCS on all trials, but the CS–UCS interval is one that has been shown not to lead to conditioning.

[1] This study was carried out as part of a project concerned with the influence of motivation on performance in learning under contract N9onr-93802, Project NR 154–107 between the State University of Iowa and the Office of Naval Research. The conditioning data were collected by Elaine Taylor.

PROCEDURE

SUBJECTS.—Sixty-four *S*s, 24 men and 40 women, from a course in introductory psychology served in the experiment. A number of other *S*s were run in the situation but were eliminated because they failed to meet certain criteria which had been established in order to exclude *S*s who made a high incidence of voluntary responses. As in an earlier experiment (**6**) the occurrence of 50% or more responses with latencies less than 300 msec. was taken as indicating a voluntary responder. Seven *S*s, four from the nonanxious group and three from the anxious group were eliminated from the experiment on this basis.

Half of the *S*s, 12 men and 20 women, constituted the anxious group. All of these *S*s scored above the eightieth percentile on the manifest anxiety scale. The nonanxious *S*s, also consisting of 12 men and 20 women, scored at the twenty-first percentile or lower on the scale.

Apparatus and method of recording.—The equipment for recording the eyelid closure and presenting the UCS was identical with that used in a previously reported study (**5**) and will not be described here.

The CS was a combination visual and auditory signal. The visual component consisted in an increase in the brightness of a 6-in. circular milk-glass disc from .05 apparent ft.-candles to 1.51 apparent ft.-candles. The auditory component was a 1000-cycle tone produced by a Hewlett-Packard oscillator. The duration of the CS on each trial during conditioning was 550 msec. with the UCS occurring 500 msec. after its onset. The duration of the UCS, an air puff of 1.0 lb./sq. in. pressure, was limited to 50 msec. by means of a 110-v., 60-cycle AC-operated solenoid valve controlled by an electronic timer. The time interval between the onset of the CS

and UCS in the extinction period was 2,500 msec.

Conditioning and extinction.—The experiment was conducted in a single period. Following the instructions each *S* received three presentations of the CS alone. A single presentation of the UCS was then administered and the response of the eyelid recorded for 40 sec. The *S* was given a ready signal, the word "ready," on each trial. He was instructed to blink once to this signal and then to look at the disc, attempting to remain as relaxed as possible. Following the thirtieth conditioning trial, *S* was reminded again to be as relaxed as possible and not to try to control his reactions to the CS or UCS.

Sixty conditioning trials were given in all with the ninth or tenth trial of each successive block of ten trials being a test trial. No UCS was administered on these trials. The intertrial interval averaged 20 sec., with intervals of 15, 20, and 25 sec. being employed in a random sequence.

A total of 40 extinction trials was given immediately following the completion of the conditioning trials. During this period the UCS was given on every trial but at an interval of 2,500 msec. following the onset of the CS. Responses occurring between 150 and 500 msec. following the CS were recorded as CR's during this period. Intertrial intervals of 20 and 25 sec. were used. At the end of the session all *Ss* were questioned as to the purpose of the experiment and warned not to discuss the experiment with other members of the class. An interesting sex difference was revealed in the ability of the *Ss* to identify correctly the experiment as being concerned with conditioning. Whereas 66.7% of the men were able to identify the experiment, only 17.5% of the women were able to do so. The experiment was conducted late in the semester some weeks after conditioning had been covered briefly in class.

Results and Discussion

Anxiety level and conditioning.—The mean numbers of CR's exhibited in the 60 conditioning trials are shown in Table 1. It will be seen that in both the male and female groups the anxious *Ss* gave a larger number of CR's than the nonanxious *Ss*. The results of an analysis of variance of these data are summarized in Table 2. It will be noted that the *F* for the anxious–nonanxious condition was 4.51, which is significant at a confidence level below .05. The *F* for sex was not significant at the .05 level but was at less than the .10 level of confidence.

TABLE 2

Analysis of Variance for Number of CR's in 60 Trials

Source	*df*	Mean Square	*F*
Anxiety	1	1251.39	4.51*
Sex	1	914.55	3.30
A × S	1	14.75	
Within groups	60	277.53	
Total	63		

* Significant at the .05 level.

Anxiety level and extinction.—Table 3 presents the data on extinction in terms of the number of CR's made in the 40 trials. It is apparent that the anxious groups show a higher level of response than the nonanxious groups. Table 4 gives the analysis of variance of these data. Once again the *F* between the anxious and nonanxious groups is significant at better than the .05 level of confidence. Unlike the data for the conditioning period, how-

TABLE 1

Mean Number of CR's in 60 Conditioning Trials

Sex	*N*	Anxious		Nonanxious	
		Mean	σ_M	Mean	σ_M
Male	12	25.1	4.24	15.0	4.43
Female	20	31.9	3.93	23.8	3.92

TABLE 3

Mean Number of CR's in 40 Extinction Trials

Sex	Anxious		Nonanxious	
	Mean	σ_M	Mean	σ_M
Male	14.7	3.03	8.42	3.23
Female	15.2	2.33	9.10	2.19

TABLE 4

ANALYSIS OF VARIANCE FOR NUMBER OF CR'S IN EXTINCTION

Source	*df*	Mean Square	*F*
Anxiety	1	606.39	5.62*
Sex	1	5.55	
A × S	1	.08	
Within groups	60	107.98	
Total	63		

* Significant at <.03 level.

ever, the extinction data show little indication of a sex difference.

Comparison with earlier study.—A comparison of the data of the present study with comparable conditioning measures from an earlier study is interesting for the degree of stability shown from one study to the other. Spence and Taylor (6) conducted an experiment in the same physical situation and under approximately the same CS–UCS time interval (500 and 520 msec.). The CS differed in the two studies in that a 1000-cycle tone was given in the present experiment in addition to the visual stimulus, which was identical with that in the first study. Two levels of the UCS were employed in the earlier study, one consisting of a pressure of approximately .6 lb./sq. in. and the other 2.0 lb./sq. in. The UCS in that instance was produced by a column of mercury falling in a manometer. In this type of system there is a rapid decrease in the magnitude of pressure and the duration is approximately 400 msec. The strength of the air pressure used in the present study was 1.0 lb./sq. in. and the pressure was maintained at a fairly constant level for 50 msec.

Table 5 provides the comparative data in terms of the number of CR's occurring in 60 conditioning trials. In both the men and the women, the anxious *Ss* are in each comparison superior in level of conditioning to the nonanxious *Ss*. On the assumption that *Ss* from the anxious and nonanxious groups are drawn from the same population one would expect, according to the binomial expansion, to obtain six differences in the same direction only 1.56% of the time. Taken in conjunction with the original study by Taylor (7) the evidence for a genuine difference in the level of response in conditioning between anxious and nonanxious *Ss* is very strong. It should be noted, however, that this study and the study by Spence and Taylor (6) were careful to control for voluntary responders. In small groups of ten *Ss* or less a disproportionate number of such *Ss* in one or the other of the comparison groups would greatly affect the difference that would be obtained.

There was a consistent sex differ-

TABLE 5

MEAN NUMBER OF CR'S IN 60 TRIALS FOR ANXIOUS AND NONANXIOUS *Ss* IN THREE EXPERIMENTS

Study	UCS in lb./sq. in.	Men			Women		
		N (Pairs)	Anxious	Nonanxious	*N* (Pairs)	Anxious	Nonanxious
Spence & Taylor (6)	.6	15	22.0	13.5	10	25.3	17.2
Present Study	1.0	15	25.1	15.3	10	30.0	22.2
Spence & Taylor (6)	2.0	12	22.1	15.0	20	30.8	23.3
Combined		42	22.9	14.5	40	30.0	22.2

ence in all three subexperiments, the women *Ss* responding at a higher frequency level than the men. Again all six appropriate comparisons come out in the same direction. This finding suggests that in these extreme groups on the anxiety scale the women show a significantly higher level of response in the conditioning situation than the men.[2] One possible interpretation of this finding is that the situation (novelty and strangeness of the experimental setup) arouses greater fear or anxiety in women than in men and hence women have a higher drive (*D*) level during the course of the experiment.

Summary

Two groups of 32 *Ss*, selected on the basis of extreme scores made on a test of manifest anxiety, were conditioned for 60 trials and then given 40 "extinction" trials in which the CS–UCS interval was increased to one which has been shown to be nonconducive to human eyelid conditioning.

The results indicated that there was a significant difference ($p < .05$) in performance level both during conditioning and "extinction" between anxious and nonanxious *Ss*. The findings also suggested that in these extreme groups on the anxiety scale women perform at a higher level than men. There was no indication of a difference between men and women in the extinction series.

(Received September 10, 1952; converted to early publication October 13, 1952)

[2] The Mann-Whitney test was used to test the hypothesis that there was no difference with respect to the distribution of scores of anxious men and women (the three sets of experimental data were combined since the differences among them were nonsignificant). The z value obtained was 1.92, which is significant at approximately the .05 level of confidence for a single-tailed hypothesis. A similar calculation for the nonanxious *Ss* indicated that the hypothesis of no difference between men and women distributions could be rejected at about the same level of confidence ($z = 1.97$).

References

1. Bitterman, M. E., & Holtzman, W. H. Development of psychiatric screening of flying personnel: III. Conditioning and extinction of the galvanic skin response in relation to clinical evidence of anxiety. USAF School of Aviation Medicine, Project No. 21-37-002, 1952.
2. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–100.
3. McAllister, W. R. The effects of the length and the shift of the length of the CS–US interval on conditioning performance. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
4. Schiff, E., Dougan, C., & Welch, L. The conditioned PGR and the EEG as indicators of anxiety. *J. abnorm. soc. Psychol.*, 1949, **44**, 549–552.
5. Spence, K. W. Learning and performance in eyelid conditioning as a function of the intensity of the UCS. *J. exp. Psychol.*, 1953, **45**, 57–63.
6. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
7. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
8. Welch, L., & Kubis, J. The effect of anxiety on the conditioning rate and stability of the PGR. *J. Psychol.*, 1947, **23**, 83–91.
9. Welch, L., & Kubis, J. Conditioned PGR (psychogalvanic response) in states of pathological anxiety. *J. nerv. ment. Dis.*, 1947, **105**, 372–381.

COMPLEX LEARNING AND CONDITIONING AS A FUNCTION OF ANXIETY[1]

I. E. FARBER AND KENNETH W. SPENCE

State University of Iowa

According to Hull's (**3**) theoretical formulation relating response strength to drive, all habit tendencies activated in a given situation are multiplied by the value of the total effective drive (*D*) then operating. In situations in which there is but a single or highly dominant response tendency, as in classical conditioning, an increase in the strength of the effective drive should, on this hypothesis, increase the probability of occurrence of the response and thus lead to better performance. However, in situations in which a number of different competing response tendencies are aroused, the effect upon performance of an increase in drive strength would depend upon the strength of the correct response tendency relative to the other incorrect tendencies. If the habit strength of the correct response should be relatively weak, an increase in drive should further increase the strength of the incorrect tendencies relative to the correct tendency, resulting in impaired performance. Furthermore, the degree of impairment should be a positive function of the number and strength of the competing incorrect response tendencies.[2]

One method of studying the effects of variations in the drive level of human *Ss* has been to select individuals on the basis of their scores on the Taylor Anxiety Scale (**10**). These scores, it has been assumed, are an index of *Ss'* state of *reactivity* or *excitability*, which in turn reflects their general level of drive (*D*). On this supposition it has been predicted and confirmed that relatively anxious *Ss*—those obtaining high scores on the anxiety scale—would exhibit a higher level of performance in classical eyelid conditioning than relatively nonanxious *Ss* (**9, 10**).[3] It has also been predicted that in complex learning tasks, i.e., those involving strong competing responses, the performance of anxious *Ss* would be impaired relative to that of nonanxious *Ss*, the degree of this impairment being positively related to the number and strength of the interfering responses. This pre-

[1] This study was carried out as part of a project on the influence of motivation on performance in learning under Contract N9onr-93802, Project NR 154–107, between the State University of Iowa and the Office of Naval Research. The data on maze learning were collected by Robert S. Dvorin and the conditioning data by Elaine Taylor.

[2] As Taylor and Spence have indicated, "even in instances in which the habit strength of the correct response is stronger than any of the alternative superthreshold responses, the effect of increasing the drive, and hence the number of superthreshold responses, would be to increase the number of errors [due to the operation of oscillation (sO_R)]" (**11**, p. 62).

[3] Hilgard, Jones, and Kaplan (**2**) did not find a significant relation between amount of anxiety as measured by the Taylor Anxiety Scale and amount of eyelid conditioning to the positive stimulus in a study of differential conditioning. It seems fair to point out, however, that they used only a relatively small number of *Ss* whose anxiety scores were comparable to those used in the studies by Taylor and by Spence and Taylor. They did obtain a highly significant difference in favor of anxious *Ss* in respect to frequency of responses to the negative stimulus (a result that would be predicted on the present hypothesis). The failure to obtain a similar result for responses to the positive stimulus is somewhat puzzling in view of the high correlation ($r = .73$) they report between the frequency of reponses to the positive and negative stimuli.

diction has also been verified (**5, 7, 11**).

The present study represents a further investigation of the performance of anxious and nonanxious *Ss* on a task involving response competition, namely the learning of a ten-choice stylus maze in which the level of difficulty of the choices had been previously established in an earlier investigation. Since most of these *Ss* also served in an eyelid-conditioning experiment, which is reported separately (**8**), an analysis is also reported in the present paper concerning the quality of performance of the anxious and nonanxious groups in the two situations.

Procedure

Subjects.—The *Ss* were selected from undergraduate psychology courses on the basis of their scores on a modified form of the Taylor Anxiety Scale. The anxious and nonanxious groups in the maze study each consisted of 40 students whose scores fell respectively within the upper and lower 20% of scores for a standardization population of about 2000 students. Of these, 28 anxious and 26 nonanxious *Ss* also served in the conditioning experiment. All *Ss* were naive with respect to the experimental tasks and were unaware of the reason for their selection.

Apparatus and procedure.—The stylus maze was of a conventional design cut in a piece of ⅛-in. brass sheeting 9¾ in. wide and 12¾ in. long. It consisted of a series of ten T choice points, whose relative levels of difficulty had been established some ten years previously in terms of the total number of errors made at each choice point by a group of 20 unselected *Ss* from an undergraduate course in experimental psychology. The sequence of correct choices and the rank order of difficulty of each, as determined by the earlier investigation were as follows (higher ranks represent greater difficulty):

Choice	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Corr. R	R	L	R	L	L	R	L	R	R	L
Difficulty	3	1	7	2	10	5	9	4	6	8

The maze rested on a metal-covered table behind a cloth screen that prevented *S* from seeing the maze. An ordinary pencil was used as the stylus.

Each *S* was instructed to move forward to the first choice point at the sound of a buzzer and then sideways, right or left. If the choice was correct, *E* said "yes"; if it was wrong, *E* said "no," in which case *S* corrected his response by moving in the opposite direction, past the choice point. When the buzzer again sounded, *S* moved up to the next choice point and made his choice. A demonstration maze differing in design from the experimental maze was placed in front of the screen during the reading of the instructions and each *S* was given a brief practice in responding to the buzzer and in correcting his response following a wrong choice. Following this practice period the demonstration maze was removed.

During the experiment the buzzer was sounded at 5-sec. intervals, so that the time allowed for a choice at each point was relatively constant for all *Ss*. At the end of each trial, *E* moved *S*'s hand back to the starting position for the next trial. The intertrial interval was approximately 10 sec. and each *S* was run to a criterion of two successive errorless trials.

After an interval of one to several days following the maze learning, most of the *Ss* served in an eyelid-conditioning experiment in which 60 conditioning trials and 40 extinction trials were given, according to a procedure described by Spence and Farber (**8**).

Results and Discussion

Stylus maze performance.—The means, medians, and *SD*'s of the number of errors and the number of trials to the criterion of maze learning for

TABLE 1

Stylus Maze Performance of Anxious and Nonanxious *Ss*

Group	*N*	Errors			Trials to Criterion		
		M	*Mdn*	*SD*	*M*	*Mdn*	*SD*
Anxious	40	19.15	17.0	10.61	8.28	8.0	2.99
Nonanxious	40	13.70	11.5	8.05	6.88	6.0	3.06

the anxious and nonanxious *Ss* are presented in Table 1. These data indicate that the nonanxious *Ss* were superior to the anxious in terms of both number of errors ($t = 2.56$, $p < .02$) and number of trials to the criterion of mastery ($t = 2.03$, $p < .05$). Since the distributions for both measures departed from normality, the differences between the groups were also analyzed by means of a nonparametric test of significance (**4**), with similar results. In the case of the error scores the x/σ value was 3.07 ($p = .002$), and for the trials measures x/σ was 2.57 ($p = .01$).

These results are highly similar to those obtained in the serial learning situation previously studied by Taylor and Spence (**11**) and are consonant with the theoretical expectation that nonanxious *Ss* would perform better than anxious *Ss* in a complex learning situation.

A further implication of the present theory was that the advantage of the nonanxious over the anxious group would be positively related to the level of difficulty of the individual choice points, on the assumption that relative difficulty (number of errors) at a choice point reflects the relative strength of competing incorrect response tendencies at that point. Accordingly, the total number of errors made by the two groups was computed for each choice point. These data are presented in Table 2. They indicate that, although the nonanxious *Ss* were superior to the anxious at every choice point but one, the extent of this superiority tended to vary with the difficulty of the choice point as determined by the performance of an unselected group (see Procedure). The rank-order correlation between the level of difficulty of a choice point and the magnitude of the difference between the number of errors by the two groups at that point was found to be .74, which for 9 *df* yields $p < .01$. This result confirmed a similar finding by Taylor and Spence.

Since the difference between the two groups tended to be greater at the more difficult choice points than at the less difficult ones, it follows that an increase in difficulty impaired the performance of the anxious *Ss* to a greater extent than that of the nonanxious *Ss*. Accordingly, it would be expected that the difference between the number of errors made at the difficult choice points and the number of errors at the less difficult choice points would be greater in the case of the anxious group. The mean difference between the number of errors made by the anxious *Ss* at the five easiest choice points—1, 2, 4, 6, 8—and the five hardest choice points—3, 5, 7, 9, 10—was 7.30, $SD = 6.21$. The corresponding mean difference for the nonanxious *Ss* was 5.10, $SD = 3.28$. A *t* test of this difference yields a value of 1.96, which for 78 *df* is significant at

TABLE 2

NUMBER OF ERRORS AT EACH CHOICE POINT FOR ANXIOUS AND NONANXIOUS *Ss*

Group	*N*	Choice Point									
		1	2	3	4	5	6	7	8	9	10
Anxious	40	32	39	66	66	124	50	133	50	83	123
Nonanxious	40	22	26	40	40	88	53	100	31	58	90

$< .03$ on a single-tailed hypothesis.[4] Thus, the results of this test also conform to theoretical expectation.

Although the nonanxious *Ss* were relatively better than the anxious *Ss* at the more difficult choice points than at the easier ones, they nevertheless made fewer errors, on the average, even at the easiest choice points. Unfortunately, it is not possible to specify independently of *Ss'* performance whether or not there were competing responses at these easy choice points. Their ease of learning indicates that little or no competition occurred. If this were so, then as Taylor and Spence have suggested in connection with similar findings in their study, such results are not entirely in accord with the theory.

Another possible interpretation of these results in both experiments is that the difference between the performance of the anxious and nonanxious groups is due, not to a difference in their drive level, but rather to a difference in their learning ability. A test of this interpretation might be made by analyzing the performance of anxious and nonanxious *Ss* whose total scores are identical, i.e., of *Ss* who made exactly the same number of errors. On the assumption that there was no difference in the drive level of these paired *Ss*, it follows that they must also have been equated with respect to learning ability, as their performances were the same. Given equal learning ability, the performance of the two groups should not be expected to differ for the easy and difficult choice points.

To satisfy the conditions of the proposed test of the foregoing interpretation, anxious *Ss* were paired with nonanxious *Ss* who made identical total scores. The mean number of errors for 20 pairs thus selected was 13.55, $SD = 4.20$. The means and *SD*'s of the scores for the two matched groups at the five most difficult and at the five least difficult choice points are presented in Table 3. It may be seen from these results that the anxious *Ss* performed better at the easier choice points and more poorly at the more difficult choice points than did the nonanxious *Ss*. The t for matched groups between the two means at the easier points was 1.84, which for 19 df yields $p < .05$, by a single-tailed test. (The test of the difference between the means of the two groups at the more difficult choice points would, of course, yield identical results in the opposite direction.) These results indicate that these particular anxious and nonanxious groups differed with respect to some factor other than learning ability alone. They are consistent, instead, with the view that they differed in drive level, the difference sometimes leading to superior performance by the former group and sometimes to superior performance by the latter group, depending on the difficulty of the choice points.

TABLE 3

ERRORS OF 20 PAIRS OF ANXIOUS AND NONANXIOUS *Ss* AT FIVE MOST DIFFICULT AND FIVE LEAST DIFFICULT CHOICE POINTS

Group	Least Difficult Choices		Most Difficult Choices	
	M	*SD*	*M*	*SD*
Anxious	3.45	1.75	10.10	3.79
Nonanxious	4.40	2.22	9.15	2.97

[4] Although the variances of these two distributions of difference scores were dissimilar, Norton (6) has recently shown that variance differences of this magnitude do not to any appreciable degree affect the results of the t test of significance.

Comparison of stylus maze performance and eyelid conditioning.—In order to compare the performance in maze learning and in conditioning of those *Ss* who served in both tasks, it was necessary to transform the scores for the two tasks into comparable units. Accordingly, the stylus maze scores (number correct) for the 26 anxious [5] and 26 nonanxious *Ss* who served in both tasks were placed in a single distribution and translated into normalized values on a linear scale of 100 points with a mean of 50, in accordance with a procedure suggested by Guilford (**1**, pp. 247–250). The conditioning scores (total number of conditioned responses for 60 conditioning and 40 extinction trials) of these 52 *Ss* were transformed in the same manner.

TABLE 4

TRANSFORMED SCORES OF 26 ANXIOUS AND 26 NONANXIOUS *Ss* IN TWO TASKS

Group	Maze Learning		Conditioning	
	M	*SD*	*M*	*SD*
Anxious	43.04	17.64	54.42	16.15
Nonanxious	56.88	17.46	45.38	20.53

The means and *SD*'s of these transformed scores for the anxious and nonanxious *Ss* in the two tasks are presented in Table 4. It may be seen from these data that the very same anxious *Ss* who were inferior to the nonanxious *Ss* in maze performance were superior to that same nonanxious group in conditioning.

TABLE 5

ANALYSIS OF VARIANCE OF TRANSFORMED SCORES OF ANXIOUS AND NONANXIOUS *Ss* IN MAZE LEARNING AND CONDITIONING

Source	*df*	*ms*	*F*
Between *Ss*			
Between groups	1	150.24	
Error (B)	50	317.74	
Within *Ss*			
Between tasks	1	.08	
Groups × Tasks	1	3404.09	9.54**
Error (W)	50	356.75	

** Significant at .01 level of confidence.

An analysis of variance based on these results is presented in Table 5. This analysis indicates that only the interaction between group and task was significant, i.e., the difference between the performance of the anxious and nonanxious *Ss* depended on the task. No main effects were found, of course. When the simple effects were considered, the anxious *Ss* were found to be inferior to the nonanxious *Ss* in maze performance ($p < .01$, $F = 7.84$, $df = 1$ and 50) and superior to the nonanxious group in conditioning ($p < .05$, $F = 3.34$, $df = 1$ and 50). The performance of the anxious *Ss* was relatively better in conditioning than in maze learning ($p < .03$, $F = 4.72$, $df = 1$ and 50), whereas the performance of the nonanxious *Ss*, on the other hand, was relatively better in maze learning than in conditioning ($p < .03$, $F = 4.82$, $df = 1$ and 50).[6]

These results are clearly in accord with the supposition that anxious and nonanxious *Ss* differ in respect to drive

[5] The number of anxious *Ss* who served in both studies was 28. Of these, 11 were men and 17 were women. Since there is evidence (**8**) that conditioning scores may be related to sex, it was considered desirable to keep the sex-ratio constant in the anxious and nonanxious groups. The latter group consisted of 11 men and 15 women. Accordingly, the data for two women who had the lowest anxiety scores in the anxious group were omitted from the present analysis. Their inclusion would not have affected the results.

[6] All probability values in the analysis of these simple effects were halved in accordance with a single-tailed hypothesis since the direction of the difference conformed in each instance to theoretical expectation.

level (*D*) rather than in any general ability to learn. Furthermore, they support the view that no blanket statement can be made concerning the relation of motivation to the quality of performance. Whether an increase in motivation (*D*) will benefit performance or impede it depends, to an important extent, upon the nature of the task involved.

Summary

The present study was concerned with the performance of anxious and nonanxious *Ss* on a stylus maze. It was predicted that the total performance of the nonanxious group would be superior to that of the anxious group in this relatively complex task, with the extent of their superiority at each choice point being a positive function of its difficulty. In addition, a comparison was made of the scores of a number of these same anxious and nonanxious *Ss* in eyelid conditioning, a task in which the anxious group was expected to show superior performance.

Two groups of 40 *Ss* each, whose scores were respectively within the upper and lower 20% of scores on the Taylor Anxiety Scale learned the maze. Later, 26 *Ss* from each group served in the conditioning experiment.

The maze performance of the anxious *Ss* was significantly poorer than that of the nonanxious *Ss*, with the more difficult points of choice providing the greatest difference between the two groups. In conditioning, however, the anxious *Ss* were superior to the nonanxious.

On the basis of these results it was concluded that the anxious and nonanxious groups in this study differed with respect to drive level (*D*) rather than general learning ability, and that the effect of variations in drive level upon performance is a function of specific characteristics of the given task.

(Received September 10, 1952; converted to early publication October 13, 1952)

References

1. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.
2. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–99.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
4. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
5. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
6. Norton, D. W. An empirical investigation of the effects of non-normality and heterogeneity upon the *F*-test of analysis of variance. Unpublished doctor's dissertation, State Univer. of Iowa, 1952.
7. Ramond, C. K. Anxiety and task as determiners of verbal performance. Unpublished master's thesis, State Univer. of Iowa, 1952.
8. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.
9. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
10. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
11. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.

OBJECT LOCATION IN A COMPLEX PERCEPTUAL FIELD[1]

CHARLES W. ERIKSEN

The Johns Hopkins University

If one is faced with the job of locating a certain kind of object from among a group of objects, there are a number of factors that can make the task easy or difficult. The present study is concerned with isolating and determining the effects of certain of the perceptual variables that are involved in such a task. As a point of departure, one may roughly distinguish two classes of variables; those that apply to the group of objects as a whole, field variables, and those that apply to the objects sought, target variables.

An important characteristic of any perceptual field is its degree of heterogeneity. In the present problem, if the objects in the entire group differ from one another along only one dimension, for example, form, the field is relatively more homogeneous than if the objects also differ in hue or size as well as in form. Various kinds of field heterogeneity can be obtained by introducing different dimensions of variation among the objects in the field.

One of the ways the target objects can be varied is in terms of the number of characteristics by which they can be identified. They may differ from all the other objects in the field on one or several dimensions. The targets may differ from the remaining objects, for example, only on form or on both form and hue. Throughout the remainder of the paper this variable will be referred to as "target definition." Multiply defined targets are those that can be identified by unique values on two or more dimensions, while singly defined targets can be identified by their value on only a single dimension.

Two experiments are reported below. The first of these determined the time required to locate the target objects for different field heterogeneities and target definition. The second experiment determined location time when target definition was held constant but field heterogeneity was varied.

Apparatus and General Procedure

In both experiments reported below a field of objects was presented to *S* by means of a display board. The display and its supplemental equipment have been fully described elsewhere (2). Briefly, the display was a square matrix ruled off into 49 4-in. squares. The objects to be presented to *S* were mounted in the center of these squares. A sliding panel hid the display from *S*. When a catch was released, the panel dropped, giving *S* a full view of the display. The dropping panel also activated an electric timer. The *S* located objects in the display by touching a stylus to a small electrode located at the bottom of each of the squares. When all objects had been correctly located, an electronic circuit automatically stopped the timer.

The objects were paper cutouts which were pasted on 3-in. white cardboard squares. Objects were varied on the four visual dimensions: form (F); hue (H); size (S); and brightness (B). On each of these dimensions seven different values were chosen. The values selected for each of the dimensions are shown in Table 1.

Six target objects and 36 nontarget objects were always placed in the display. This left seven blank squares randomly distributed throughout the board. The location of the six target objects in the display was in accord with

[1] This research was done under Contract N5-ori-166, Task Order I, between the Systems Coordination Division, Naval Research Laboratory, Office of Naval Research, and The Johns Hopkins University. This is report No. 166-I-161, Project Designation No. NR 507-470, under that contract.

TABLE 1

THE SEVEN OBJECT VALUES USED WITHIN THE FOUR DIMENSIONS

Values	Dimensions			
	Hue*	Form	Brightness*	Size (In.)
1	R 5/6	Circles	N 8/	4/8
2	YR 5/6	Hexagons	N 7/	5/8
3	Y 5/6	Diamonds	N 6/	6/8
4	GY 5/6	Triangles	N 5/	7/8
5	G 5/6	Crosses	N 4/	8/8
6	BG 5/6	Stars	N 3/	9/8
7	B 5/6	Squares	N 1/	10/8

* Munsell notation.

eight standard patterns that had been determined empirically to be of equal difficulty.

The *Ss* were all male undergraduate students at The Johns Hopkins University. All were screened with a shortened version of the Ishihara Color Vision Test (1) to insure that they were color normal.

EXPERIMENT I

Purpose.—In a study previously reported (2) it was found that a target group which was identifiable by only one characteristic could be located at least as rapidly from a relatively homogeneous field as could a multiply defined target from a heterogeneous field. For example, when the field varied only in hue and the target objects were defined as a specific hue, location was at least as rapid as when the field varied in hue, size, and brightness and the target was defined on all three of these dimensions.

In this prior study *Ss* were not informed of the specific definition of the target group. Instead they were shown a sample of the target and told to locate the six objects in the field that were identical with the sample. It is possible that the relatively slow location obtained for multiply defined targets may have resulted from a greater complexity of judgment required of *S* by these instructions. The *S*, in selecting an object, may have had to make sure that it was not only the right hue, but also the right size and brightness as well.

The present study is essentially a repetition of the previous study, but using different instructions to the *Ss*. Here *Ss* were made explicitly aware of the complete definition of the target objects. When multiply defined targets were used, *Ss* were informed as to each characteristic by which targets could be identified.

Design and Procedure.—Table 2 shows the various field heterogeneities and target definitions used in this experiment. It should be noted that the target definition varies with the field heterogeneity. If the field involves two dimensions, the target objects can be identified by characteristic values on these two dimensions.

TABLE 2

TARGET AND NONTARGET OBJECTS USED WITH THE VARIOUS FIELD HETEROGENEITIES

Field Heterogeneity	Target Definition	Nontarget Objects
H	Red (R 5/6) squares	Differed from target on remaining six hues.
F	Gray circles	Differed from target on remaining six forms.
HF	Red (R 5/6) circles	Differed from target on combinations of the remaining six forms and hues.
HS	Red (R 5/6) 4/8-in. squares	Differed from target on combinations of the remaining six hues and sizes.
HB	Red (R 8/6) squares	Differed from target on combinations of the remaining six hues and brightnesses.
FS	Gray 4/8-in. circles	Differed from target on combinations of the remaining six forms and sizes.

TABLE 2—*Continued*

Field Heterogeneity	Target Definition	Nontarget Objects
FB	Circles with brightness N 8/	Differed from target on combinations of the remaining six forms and brightnesses.
HFS	Red (R 5/6) 4/8-in. circles	Differed from target on combinations of the remaining six hues, forms and sizes.
HFB	Red (R 8/6) circles	Differed from target on combinations of the remaining six hues, forms and brightnesses.
HSB	Red (R 8/6) 4/8-in. squares	Differed from target on combinations of the remaining six hues, sizes and brightnesses.
FSB	Circles 4/8-in. with brightness N 8/	Differed from target on combinations of the remaining six forms, sizes and brightnesses.
HFSB	Red (R 8/6) 4/8-in. circles	Differed from target on combinations of the remaining six hues, forms, sizes and brightnesses.

If three dimensions are involved, the target objects can be identified by characteristic values on all three.

Two experimental groups were involved. In the Hue Group, the *S*s were given instructions which favored a set for hue differences. As an example, instructions for locating the target objects for the field heterogeneity HFB were: "The objects in the display are of different colors, shapes, and brightness. Locate the six red objects. The red objects are the only *circles* in the display and they are also the *lightest* objects. Locate the six red objects." The same instructions were used for the target objects for the other field heterogeneities with the appropriate word substitutions. Care was taken to insure that the *S* completely understood the instructions. Speed in performance was stressed but the *S* was also cautioned against making errors. The eight field heterogeneities H, HF, HS, HB, HFS, HSB, and HFSB were used in this group.

The Form Group differed from the Hue Group in that the instructions favored a set for form differences. The instructions were of the same order as for the Hue Group with the exception that they began and ended with "Locate the six circles" instead of "Locate the six red objects." This group contained the heterogeneities F, HF, FS, FB, HFS, HFB, FSB, and HFSB.

Eight *S*s were used in each group. Within groups, each *S* was tested once on each field heterogeneity. The order of testing on the different field heterogeneities within groups was systematically varied between *S*s.

Results

All location times were transformed into two-place logarithms and a separate analysis of variance was performed on the data for each of the two groups. In both groups, field heterogeneities with their associated singly or multiply defined targets constituted a significant source of variance ($p < .001$). Table 3 presents the mean location time (in logarithms) obtained with the different field heterogeneities and target definitions.

TABLE 3

MEANS FOR THE FIELD HETEROGENEITIES AND THEIR ASSOCIATED TARGET DEFINITIONS FOR THE HUE AND FORM GROUPS*

Hue Group		Form Group	
Field Heterogeneity	Mean	Field Heterogeneity	Mean
HF	.70	HF	.71
H	.77	HFB	.83
HFB	.77	FB	.87
HFS	.78	HFS	.87
HB	.78	F	.89
HS	.79	HFBS	.89
HBS	.81	FS	1.00
HFBS	.85	FBS	1.04

* The means in the table are the means of the logarithms of location time in seconds. Within groups, mean differences of .07 and .09 are significant at approximately the .05 and .01 levels. In comparing means between groups, differences of .13 and .18 are significant at approximately the .05 and .01 levels. The magnitude of differences needed for significance at the .05 and .01 levels was computed by pooling the sums of squares for the respective error terms for both groups and for Exp. II.

Considering first the Hue Group, we find that target objects defined only on the hue dimension can be located in a relatively homogeneous field, varying only in hue, as rapidly or more rapidly than multiply defined targets can be located from more heterogeneous fields. The only exception is when the field is heterogeneous in hue and form and the target is defined on both of these dimensions. In this case there is some indication that heterogeneity and definition on both hue and form is superior to heterogeneity and definition on the hue dimension alone.

Similar results are found for the Form Group. Here again heterogeneity in hue and form produces the most rapid location, but the times for the other heterogeneities are not significantly faster than those obtained for form heterogeneity. In fact, targets defined on both form and size, or form, size, and brightness, when located in fields that are heterogeneous on these dimensions, yield significantly slower location than that obtained for form definition and heterogeneity.

The results from both the Form and Hue Groups support the previously reported data (**2**). Here the results indicate that even when *S*s are explicitly aware that the target objects can be identified on each of the defining dimensions, multiply defined targets in heterogeneous fields have no consistent advantage in location time over singly defined targets in more homogeneous fields.

A question may be asked whether the differential sets favoring form or hue differences affected location times. In comparing the field heterogeneities HF, HFS, HFB, and HFSB, which occurred in both groups and with the same target definitions, it is found that the location for all four was faster in the Hue Group. However, when the differences between the two groups were evaluated via a modified three-way analysis of variance (groups, field heterogeneities, and individuals), neither the *F* test for between groups nor the *F* test for interaction between groups attained significance at the .05 level. Thus there is no reliable evidence that performance was differentially affected by the form or hue set.

Experiment II

Purpose.—In Exp. I the two variables, field heterogeneity and target definition, have been confounded. When the heterogeneity of the field was varied on one, two, three, or four dimensions, the target definition also varied on the same dimensions. Consequently, it is impossible to separate the effects of the respective variables. The location time for targets defined on two or more dimensions cannot be directly compared with targets defined on only one dimension since the field heterogeneities also differed. The present experiment was performed in order to permit a comparison between multiply defined targets and a singly defined target under conditions of comparable field heterogeneity.

Design and Procedure.—Since location times were obtained for multiply defined targets in Exp. I, the present experiment need only determine the location times for a singly defined target in fields of comparable heterogeneity. Multiple and single target definition can then be compared by comparison of the data from the two experiments.

Seven of the eight field heterogeneities used in the Form Group of Exp. I were also used in this experiment. They were HF, FS, FB, HFS, HFB, FSB, and HFSB. The target objects were defined only on the form dimension. For all seven field heterogeneities, the target objects were six circles. For each field heterogeneity the target objects, as a group, could be distinguished from the remaining 36 objects in the field only on the basis of this form difference.

The procedure was the same as in Exp. I, with the exception of the instructions and infor-

mation given the *Ss*. The general nature of the instructions in this experiment can be seen in the specific instructions used when the field was heterogeneous in HFB. "The objects in the display are of different forms, colors and brightnesses. Locate the six circles."

An additional seven *Ss* were used. Each *S* was tested once on each of the seven field heterogeneities with the order of testing systematically varied between *Ss*.

Results

Location times were again transformed into two-place logarithms. A two-way analysis of variance (field heterogeneity and individuals) was performed on the transformed data. The *F* test for differences in field heterogeneity was significant beyond the .01 level. The mean location times (in logarithms) for the seven field heterogeneities are shown in Table 4.

TABLE 4

Means for Singly Defined Targets*

Field Heterogeneity	Mean
HF	.91
HFB	.94
FS	1.00
FB	1.01
HFS	1.03
HFBS	1.06
FBS	1.07

* The means in the table are the means of the logarithms of location time in seconds. Mean differences of .07 and .09 are significant at the .05 and .01 levels. In comparing the means in the above table with means in Table 2, mean differences of .13 and .18 are significant at the .05 and .01 levels.

Examination of the values in this table show that in nearly every case, location time is appreciably longer than was obtained for the same field heterogeneities in Exp. I. A precise comparison can be had by comparing the times for the seven field heterogeneities in this experiment with the corresponding ones in the Form Group of Exp. I by means of a modified three-way analysis of variance (field heterogeneity, singly vs. multiply defined target, and individuals). This analysis yields significant *F* tests ($p < .01$) for between-field heterogeneities, between singly and multiply defined targets, and for the interaction between field heterogeneity and singly vs. multiply defined targets.

The significant *F* test for field heterogeneity is only a confirmation of the previous analyses for differences between heterogeneities. The significant *F* test for between singly and multiply defined targets, however, indicates that with comparable field heterogeneity, multiply defined targets result in faster location.

The interaction between field heterogeneity and singly vs. multiply defined targets can be accounted for, at least in part, by considering the rank-order correspondence between the different fields in the two experiments. In both experiments, heterogeneity in HF results in the most rapid location while heterogeneity in FBS gives the slowest. With the exception of heterogeneity in FS, the rank-order correspondence is perfect. The discrepancy for FS can account for the obtained interaction. While multiply defined targets are located more rapidly under most field heterogeneities, for FS heterogeneity, location time is the same for both singly and multiply defined targets.

A possible explanation for this may be found in the dependence of the apparent size of an object upon the object's form. This interaction between apparent size and form coupled with the fact that the size differences involved were quite small, may have resulted in size differences being a useless cue to the *S* in the task of location. If this was the case, the *S*'s task was essentially the same in both experiments. In both, the field varied in form and size and *S* had to rely

solely upon form differences for successful location of the target objects.

Discussion

The point of departure for the present study concerned the factors that influenced the ease with which designated objects could be located from among a large group of objects. The results of the two experiments provide some answers to the problem. These results establish the importance and suggest the role that the variables of field heterogeneity and target definition play in determining speed of location. Locatability depends upon the kind of heterogeneity that exists among the objects in the group and also upon the number of ways the target objects can be defined or identified.

The role of field heterogeneity is not clear, partly as a result of the inability to quantify or order this variable. We cannot say that a group of objects that differ from one another in hue and form is more or less heterogeneous than a group that differs in form and size. However, there are clear differences in location times for different kinds of field heterogeneity. The results from both experiments, I and II, are consistent in showing that when the objects vary on hue, hue and form, or hue, form, and brightness, location time is significantly faster than when the objects differ from each other on form, size, and brightness or on hue, form, size, and brightness. There is some tendency for heterogeneities involving the hue dimension to result in faster location, but this may well be an artifact of the discriminability between the adjacent values that were chosen for the four dimensions.

The faster location obtained for multiply defined as opposed to singly defined targets is somewhat limited by the comparisons that were involved in the present study. The results do show that under conditions of comparable field heterogeneity, target objects that are uniquely defined on each of the dimensions that vary in the field are located more rapidly than those that are defined on only the form dimension. Thus if the field is heterogeneous in terms of form and hue, target objects that can be identified by either their form or their hue are more rapidly located than are objects identifiable only by their form. It seems likely that similar results would be obtained if multiply defined targets were compared with targets that were singly defined on the hue, size, or brightness dimensions. Such an outcome might be expected, not only from the above results, but also from other experiments (**3, 4**) which have indicated that compounding stimulus differences improves discriminability.

Summary

The present study was concerned with the perceptual variables that determine the speed with which designated objects can be located from among a large field of objects. Two variables, field heterogeneity and target definition, were studied.

Field heterogeneity, as determined by the number of ways that the objects in the field differed from one another, was found to be an important determiner of location time. Location was quickest when the fields varied on hue, hue and form, and hue, form, and brightness, and slowest when the field was heterogeneous on hue, form, size and brightness, and form, brightness, and size.

Under comparable field heterogeneity, target objects that were multiply defined by assigning them unique

values on each of several dimensions were located more rapidly than objects defined on only one dimension.

(Received for early publication September 25, 1952; revision received November 28, 1952)

References

1. Chapanis, A. A comparative study of five tests of color vision. *J. opt. Soc. Amer.*, 1948, **38**, 626–649.
2. Eriksen, C. W. Location of objects in a visual display as a function of the number of dimensions on which the objects differ. *J. exp. Psychol.*, 1952, **44**, 56–60.
3. Grant, D. A. Perceptual versus analytical responses to the number concept of a a Weigl-type card sorting test. *J. exp. Psychol.*, 1951, **41**, 23–29.
4. Weitz, J. The coding of airplane control knobs. In P. M. Fitts (Ed.), *Psychological research on equipment design.* Washington, D. C.: U. S. Govt. Print. Off., 1947. Pp. 187–198.

QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Founded in 1948

Published Quarterly in February, May, August, and November

for the

EXPERIMENTAL PSYCHOLOGY GROUP

Editor: D. Russell Davis, University of Cambridge

Subscriptions from the United States
should be entered now for the next year.
30 shillings per year (current rate $4.20).

Order from

W. Heffer & Sons, Ltd.
Cambridge, England